

论 著

基于深度学习构建超声与胸部CT融合模型以提高对乳腺肿块的诊断能力

刘晓玲^{1,3} 肖为瀚¹ 覃夏川^{2,*}

1.首都医科大学附属北京安贞医院南充医院·南充市中心医院超声科
(四川 南充 637000)

2.成都市第二人民医院超声科
(四川 成都 610000)

3.安徽医科大学第一附属医院超声科
(安徽 合肥 230022)

【摘要】目的 本研究旨在利用深度学习中的卷积神经网络(convolutional neural networks, CNN)技术,对超声和胸部CT图像中的良性和恶性乳腺肿块进行鉴别。方法 我们回顾性地收集了281张良性肿块图像和155张恶性肿块图像,所有病例均经过手术切除和麦默通活检以获得病理结果。采用ResNet-50的CNN架构,并集成了卷积块注意力模块(convolutional block attention module, CBAM),分别构建了超声(Ultrasound, US)、CT和US+CT的深度学习模型,对这些病变的良恶性进行了分析。我们计算了模型的灵敏度、特异性、准确度以及受试者工作特征曲线下面积(Area Under the Curve, AUC),并与人工阅片结果进行了比较。结果 超声科医生阅读超声图像得到的AUC为0.808,而CT医生阅读CT图像得到的AUC为0.749。单独的超声模型的AUC为0.883,单独的CT模型的AUC为0.871,而超声加CT联合模型的AUC达到了0.947。与单独使用超声或CT进行诊断相比,超声和胸部CT的联合使用展现出了更优的诊断性能。结论 采用CNN进行的超声加CT图像深度学习分析具有很高的诊断性能,能够有效地区分乳腺超声图像上的良性和恶性肿块,其诊断效能优于单一的影像学方法。

【关键词】深度学习; 回顾性研究; 乳腺肿瘤; CT; 超声
【中图分类号】R737.9
【文献标识码】A
DOI:10.3969/j.issn.1672-5131.2025.07.028

Development of A Deep Learning-based Multimodal Model Integrating Ultrasound and Chest CT for Improved Diagnosis of Breast Masses

LIU Xiao-ling^{1,3}, XIAO Wei-han¹, QIN Xia-chuan^{2,*}.

1.Department of Ultrasound, Beijing Anzhen Nanchong Hospital of Capital Medical University & Nanchong Central Hospital, Nanchong 637000, Sichuan Province, China
2.Department of Ultrasound, Chengdu Second People's Hospital, Chengdu 610000, Sichuan Province, China
3.Department of Ultrasound, the First Affiliated Hospital of Anhui Medical University, Hefei 230022, Anhui Province, China

ABSTRACT

Objective To implement deep learning through convolutional neural networks (CNNs) to distinguish benign and malignant breast masses on US and chest CT images. **Methods** We retrospectively collected 281 images of benign masses and 155 images of malignant masses from patients who underwent surgical resection and mammotome biopsy to obtain their pathological results. The ResNet-50 architecture was used, and the convolutional block attention module was included to establish deep learning models. The benign and malignant masses were analyzed using US, CT, and US + CT images. The sensitivity, specificity, accuracy, and AUC were calculated. Moreover, the deep learning results were compared with those of manual evaluations. **Results** The AUC of the ROC curve from ultrasound doctor evaluations on US images was 0.808 and that on CT images was 0.749. In addition, the AUC of the ROC curve from the single US deep learning model was 0.883, while that of the single CT model was 0.871, and that of the combined US + CT model was 0.947. Compared with single-modality diagnosis and manual evaluations, US combined with chest CT shows the highest diagnostic performance. **Conclusion** Deep learning analysis of US + CT images using a CNN shows high diagnostic performance and effectively distinguishes benign and malignant breast masses. Its diagnostic accuracy notably outperforms the use of a single imaging modality.

Keywords: Deep Learning; Retrospective Studies; Breast Neoplasms; CT; Ultrasound

乳腺癌是全球女性中最常见的癌症类型,2018年,在女性患者中约850万新诊断的癌症中,乳腺癌占25%以上^[1]。2004年以来,乳腺癌新发病例一直保持着每年约0.3%的增长速度。因此早期发现并正确诊断乳腺癌,能有效地降低其死亡率^[2]。

超声作为一种灵活的绿色成像方式,以其无电离辐射、操作方便、成本低、成像实时显示等独特优势,已成为乳腺诊断的一线诊断手段^[3-4],但是乳房超声结果的准确率常常依赖于操作者的熟练度,同时并不能获得满意的结果^[5]。为了获得更加准确的结果,有研究将X线和B超联合诊断乳腺病变,取得比单一的影像学检查更有效的结果^[6]。

由于目前肺部结节的高发,低剂量薄层胸部CT被当成一种常见的检查手段筛查肺部病变^[7],在显示肺的同时,整个乳房组织也会出现在胸部CT上^[8-9],因为它使用更薄的切片厚度和更小的视野覆盖,比X线具有更高的高空间分辨率成像^[10]。我们认为胸部多层螺旋CT能够提供比X线更多乳腺肿瘤的信息^[11],但是常常被作为一项附加的检查方式被忽视。

人工智能因其在图像识别任务中的出色表现而受到广泛关注,其能有效提高医学影像判读的诊断准确性,同时增加诊断的客观性^[3,12]。虽然有研究使用深度学习对单一的影像手段诊断乳腺结节的良恶性^[13-17],由于单一的影像检查并不能得到满意的结果,为了更加准确客观地诊断乳腺肿瘤的良恶性,我们尝试使用深度学习建立B超与胸部CT联合模型,以提高乳腺良恶性病灶的诊断率,据我们所知这是首次使用深度学习对CT乳腺病变进行诊断,也是首次采取两种影像结合建立AI模型,同时加入了CBAM提高深度学习注意力,区分良恶性乳腺肿块的诊断性能,并通过将其与放射科医师读数进行比较来评估其诊断性能。

1 资料与方法

1.1 一般资料 该研究得到了机构审查委员会的批准,并免除了知情同意的要求。我们回顾性分析了2020年3月至2021年6月在安徽医科大学第一附属医院和南充市中心医院接受胸部CT和乳腺超声的436名女性患者,平均年龄46±15岁(年龄范围:16-82岁)。436名患者包含436个结节,其中恶性结节155个,良性结节281个。

患者在一周内进行超声及CT检查,通过手术切除和麦默通活检获得病理结果。病理结果分为良性病变和恶性病变。良性病变包括纤维瘤、导管内乳头状瘤、良性叶状肿瘤、腺样体增生、炎性肿块。恶性病变包括浸润性导管癌、导管内乳头状癌、原位癌和浸润性小叶癌。工作流程见图1,包括纳入和排除标准。

【第一作者】刘晓玲,女,主治医师,主要研究方向:超声造影、介入超声及人工智能超声的应用。E-mail: lxlzlo@126.com

肖为瀚,男,住院医师,主要研究方向:人工智能超声的应用。E-mail: 515146462@qq.com

96 • 【通讯作者】覃夏川,男,主任医师,主要研究方向:超声造影、介入超声及人工智能超声的应用。E-mail: 11326636@qq.com

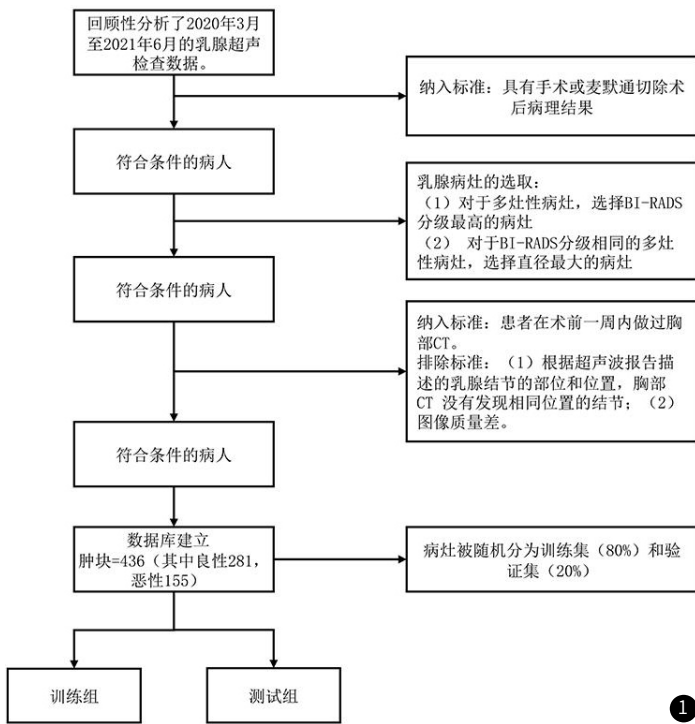


图1 患者数据筛选。

1.2 仪器与方法 本研究中使用超声仪器包括: GE LOGIQ E9(美国通用电气公司)、迈瑞Resona 7(深圳迈瑞生物医疗电子股份有限公司, 中国)、Esaote My Lab(意大利Esaote公司)。

CT图像由不同的CT扫描仪获得, 包括LightSpeed VCT(GE医疗系统, 美国威斯康星州密尔沃基市)和Philips Ingenuity 64层CT(飞利浦医疗保健, 美国俄亥俄州克利夫兰市)。

一位经验丰富的放射科医师(15年超声工作经验)在超声图像上标注了良性和恶性结节的位置。以乳头为中心描述包块的方位及距离乳头的距离, 根据超声描述, 我们在CT上以乳头为中心定

位在定位包块, 若CT上未发现病灶或者病灶在超声和CT的位置和大小不一致, 我们选择将该病例剔除。

1.3 图像预处理 在将图像引入CNN前, 我们人工消除不相关的信息, 如文本和仪器设置。首先将正方形分割掩模的尺寸调整到最大, 以保持足够的边缘(边缘定义为病变边界和分割掩模本身边界之间的距离), 同时不超过采样盒边界。然后, 由经验丰富的放射科医生进一步改变每个分割掩模, 以确保在每种成像模式中相似的病变与掩模比例。最后, 将所有截取的超声及CT图像调整为 300×300 的宽高比进行质量控制, 并作为网络输入。在深度学习中, 收集到的数据集集中的病灶被随机分为训练集(80%)和验证集(20%)。

1.4 深度学习模型开发 ResNetv50已被证明是一种具有良好稳定性和性能的特征提取结构, 包括医学图像分类^[18]。在之前的研究中, ResNetv50 的乳房模型不管在X线诊断还是超声都取得了不错的成绩^[12,19-20]。ResNet50 可以在不分离前景和背景的情况下, 通过设置锚点和边界框来检测病变并进行分类。并且能直接从病灶区域提取特征, 然后合并多个特征对病灶进行分类^[21]。

我们的模型保留了其结构, 并进行了一些调整以适合我们的数据。通过旋转、水平或垂直翻转、随机裁剪和随机通道移位等图像处理来生成更多的训练图像, 可以有效地模仿现实世界中观察到的数据多样性, 避免模型过度拟合。

由于提供数据集较小, 为了更好地进行分类, 此次采用迁移学习并微调。使用预先训练的模型进行特征提取: 利用在相同域中的较大数据集上训练的模型中学习的特征。这是通过实例化预训练的模型并在顶部添加完全连接的分类器来完成的。预先训练的模型是“冻结的”, 训练过程中仅更新分类器的权重。在这种情况下, 卷积基础提取了与每个图像关联的所有特征, 刚刚训练了一个分类器, 该分类器根据给定的提取特征集确定图像类。

为了能够获得最后注意力图和热力图, 在resnet50中加入了CBAM^[22], CBAM是一个简单但有效的注意力模块, 给定一个中间特征图, 我们沿着空间和通道两个维度依次推断出注意力权重, 然后与原特征图相乘来对特征进行自适应调整。由于CBAM是一个轻量级的通用模块, 它可以无缝地集成到任何CNN架构中, 额外开销忽略不计, 并且可以与基本CNN一起进行端到端的训练。同时他能查看模型如何识别感兴趣的区域以及模型区分癌症类别的程度。CNN架构如图2所示。

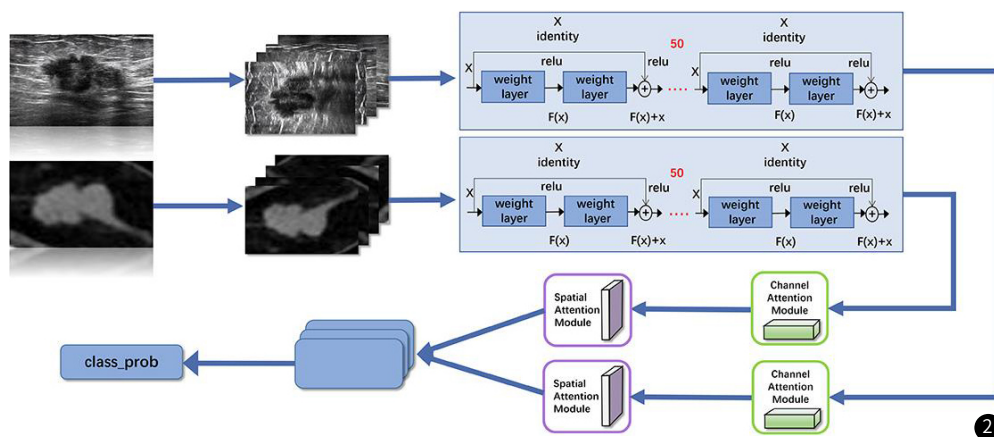


图2 深度学习框架示意图。

1.5 放射医师阅片 超声检查由两名超声科专家评估了测试组的超声图像, 他们两名不知道病理和临床结果。如果患者的超声结果存在差异, 由两人讨论后做出评估。

CT检查由两名放射科专家评估测试数据的CT图像, 他们两名不知道病理和临床结果。如果患者的CT结果存在差异, 由两人讨论后做出评估。

1.6 评估 每位医生读取的样本与AI模型一致, 该模型使用了测试集中的所有患者病例。测试数据集用于计算区分良性肿块和恶性肿块的灵敏度、特异性和准确性。进行接收器操作特征(ROC)分析以计算接收器操作特征曲线(AUC)下的面积, 用于三种CNN模

型和放射科医生阅读器区分良性肿块和恶性肿块的性能。

2 结果

超声的Loss函数, Loss在训练过程中, 1-20轮下降快速, 但是下降过程有震荡, 20-30轮小幅下降, 约30轮Loss基本不下降, Loss值反复震荡。CT的Loss值在前30轮的训练过程中下降, 略微有震荡, 在30轮之后Loss值也是反复震荡, 但最终值比超声值略微小。Merge模型在前20轮稳步下降, 震荡程度比前面两者显著降低, 说明模型在稳定收敛, 20轮之后Loss值几乎不变, 无明显震荡(图3)。

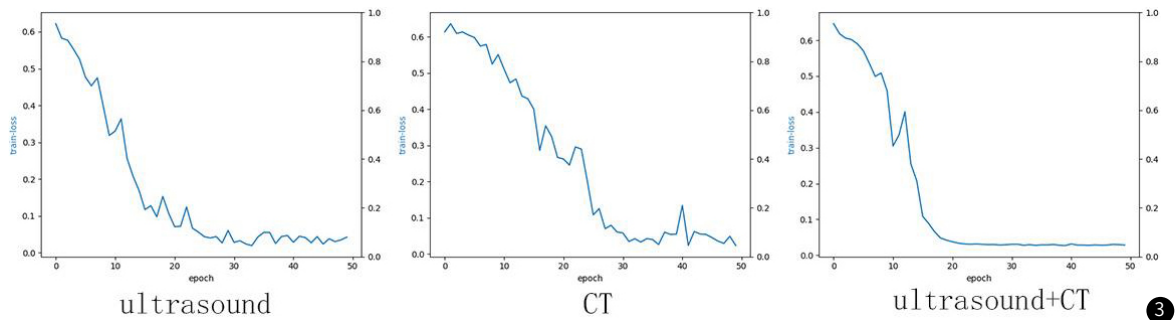


图3 三个CNN模型的Loss值。

我们的AI系统涉及三个深度学习模型，即超声模型，CT模型和超声+CT模型，为了能有效地比较三者之间的差异，我们只选择了Restnet50进行测试。单一超声的敏感性0.76，特异性0.89，准确度85.9%。单一CT的敏感性0.72，特异性0.91，准确度85%。超声+CT联合的敏感性0.83，特异性0.90，准确度88%。超声科医师阅读超声图像后诊断肿瘤的敏感性0.88，特异性0.74，准确度78.1%。放射科医师阅读CT图像后诊断肿瘤的敏感性0.92，特异性0.58，准确度68.3%(表1)。

超声科读片超声图像AUC的ROC为0.808，CT读片CT图像的ROC为0.749，单一超声模型AUC的ROC为0.883，单一CT模型AUC的ROC为0.871(图4)，超声+CT联合模型为0.947(图5)，我们发现机器阅片结果高于人工，同时，CT+超声联合模型效能高于单一超声模型和单一CT模型。

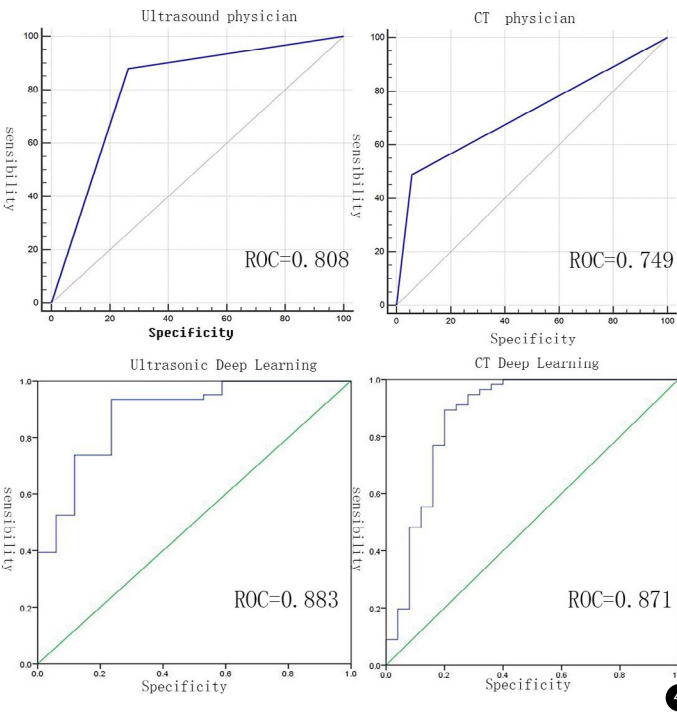


表1 人工智能诊断模型与放射科医师诊断性能的比较

	AUC	准确度	敏感性	特异性
CT 医师	0.749	0.683	0.92	0.58
US 医师	0.871	0.781	0.88	0.74
US CNN	0.883	0.859	0.76	0.89
CT CNN	0.871	0.85	0.72	0.91
US+CT CNN	0.947	0.88	0.83	0.90

CNN =deep convolutional neural network

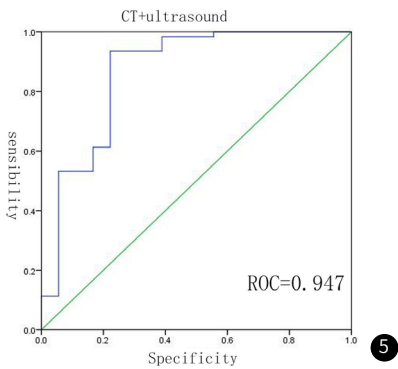


图4 放射科医生和CNN的ROC曲线对比。图5 US+CT CNN的ROC曲线。

3 讨论

超声已广泛用于术前检查乳房病变^[23]，BI-RADS词典标准化了如何描述术语并相应地对乳房肿块进行分类^[24]。然而，良性和恶性肿瘤的形态特征有时会重叠。诊断结果在很大程度上取决于诊断医师的经验，从而导致误诊。深度学习可以自动从数据中提取本质和有效的特征，并且利用大量训练数据对CNNs进行优化和训练以提高性能。为了更加客观地诊断乳腺肿瘤，给予临床医师提供更多有价值信息，我们引入了深度学习诊断乳腺病变。

一些研究者尝试通过深度学习来判断乳腺的良恶性，并且取得了可喜的成绩。Nan Zhang等收集1121名患者，使用ResNet50对测试集中202个病灶进行诊断，其准确率为77.0%，

AUC为0.846，稍优于放射科医师诊断^[12]。Jianming Li使用超声科读片超声图像的ROC为0.808，CT读片CT图像的ROC为0.749，单一超声模型AUC的ROC为0.883，单一CT模型AUC的ROC为0.871(图4)，超声+CT联合模型为0.947(图5)，我们发现机器阅片结果高于人工，同时CT+超声联合模型效能高于单一超声模型和单一CT模型。Res Net50建立模型，其在诊断良性和恶性乳腺的准确率为86.82%，AUC高达0.943^[14]。Mehedi Masud等使用ResNet50对乳腺良恶性进行诊断，其准确率为92.4%^[25]。ResNet随着层数的增加很容易优化，并且随着深度的增加其准确率增加。迁移学习还进一步提高了分类性能。在我们的研究中，

我们建立ResNet50应用在超声诊断模型，其诊断乳腺肿瘤的良恶性，AUC为0.883，优于人工诊断，与之前的研究一致。

由于目前肺部结节的泛滥，胸部的低剂量被患者接受作为筛查的一部分。由于胸部CT常常包含了乳腺的信息，胸部CT作为一项附加的诊断乳腺的工具被应用，所以我们选择胸部CT进行研究，其常被作为入院的常规筛查方法。目前，深度学习被广泛应用于肺部疾病的诊断^[26-28]，近期，Jieun Koh尝试用深度学习检测胸部CT上乳腺癌的检出，发现胸部CT能检测到96.5%的乳腺癌^[29]，其认为深度学习算法可灵敏检测胸部CT上的乳腺癌，但据我们所知，目前并没有研究使用深度学习应用在CT诊断乳腺肿瘤的良恶性。我们尝试使用CT与超声结合来判断乳腺病灶的良恶性，当使用单一CT深度学习模型诊断乳腺包块良恶性，其准确度接近超声诊断结果。进一步将超声和CT结合，可以惊喜地发现其AUC明显高于单一模态。我们认为可能两种影像手段的联合，给予机器提供了更多的信息。

虽然神经网络能够分辨乳腺肿瘤的良恶性，但是如果机器不能解释它是如何知道的，会给医生和他们的病人带来严重的困境，无法解释的神经网络系统导致其在医学成像中的应用通常被称为“黑匣子”^[30]。我们加入了一个简单但有效的注意力模块CBAM，这也是首次应用该模块对乳腺病变进行分类，其在通道注意力机制中引入全连接，并通过全连接降维，有利于提取更重要的信息(相当于PCA操作)。一般注意力机制是用AvgPool操作，但是CBAM引入MaxPool操作，可以提取到不同信息，增强特征的多样性，进而提高分类和检测的正确率。同时其提供可视化热图，使诊断医师能够查看模型如何识别感兴趣的区域以及模型区分乳腺病变类别的程度。

当然，我们研究还存在一定的局限性，首先这是一项回顾性的研究，并且病例数量有限。再者，我们提前筛选出CT和超声同时存在的乳腺肿瘤，存在一定的选择偏移。最后我们单一的使用ResNet-50对三者进行比较，与单个网络相比，不同神经网络的集成可能具有更高的准确性。

使用深度CNN模型在诊断乳腺病理良恶性取得了比人工诊断更好的结果，不管是CT和超声，进一步将超声+CT结合建立深度学习模型，进行图像分析显示出很高的诊断性能，能有效区分乳腺超声上的良恶性肿块，其诊断效能优于单一的影像方式。

参考文献

- [1] Bray, F., J. Ferlay, I. Soerjomataram, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. *CA Cancer J Clin*, 2018, 68(6): 394-424.
- [2] DeSantis, C. E., J. Ma, M. M. Gaudet, et al. Breast cancer statistics, 2019[J]. *CA Cancer J Clin*, 2019, 69(6): 438-451.
- [3] Shen, Y. T., L. Chen, W. W. Yue, et al. Artificial intelligence in ultrasound[J]. *Eur J Radiol*, 2021, 139: 109717.
- [4] Kim, G. R., J. S. Choi, B. K. Han, et al. Preoperative axillary US in early-stage breast cancer: potential to prevent unnecessary axillary lymph node dissection[J]. *Radiology*, 2018, 288(1): 55-63.
- [5] Zhang, E., S. Seiler, M. Chen, et al. BIRADS features-oriented semi-supervised deep learning for breast ultrasound computer-aided diagnosis[J]. *Phys Med Biol*, 2020, 65(12): 125005.
- [6] Houssami, N. and S. Ciatto, Ultrasound and mammography for breast cancer screening[J]. *Jama*, 2008, 300(13): 1514-1515; author reply 15.
- [7] Lim, K. P., H. Marshall, M. Tammemägi, et al. Protocol and Rationale for the International Lung Screening Trial[J]. *Ann Am Thorac Soc*, 2020, 17(4): 503-512.
- [8] Son, J. H., H. K. Jung, J. W. Song, et al. Incidentally detected breast lesions on chest CT with US correlation: a pictorial essay[J]. *Diagn Interv Radiol*, 2016, 22(6): 514-518.

- [9] Bin Saeedan, M., M. Mobara, M. A. Arafah, et al. Breast lesions on chest computed tomography: pictorial review with mammography and ultrasound correlation[J]. *Curr Probl Diagn Radiol*, 2015, 44(2): 144-154.
- [10] Monzawa, S., T. Washio, R. Yasuoka, et al. Incidental detection of clinically unexpected breast lesions by computed tomography[J]. *Acta Radiol*, 2013, 54(4): 374-379.
- [11] Boone, J. M., A. L. Kwan, K. Yang, et al. Computed tomography for imaging the breast[J]. *J Mammary Gland Biol Neoplasia*, 2006, 11(2): 103-111.
- [12] Zhang, N., X. T. Li, L. Ma, et al. Application of deep learning to establish a diagnostic model of breast lesions using two-dimensional grayscale ultrasound imaging[J]. *Clin Imaging*, 2021, 79: 56-63.
- [13] Yu, T. F., W. He, C. G. Gan, et al. Deep learning applied to two-dimensional color Doppler flow imaging ultrasound images significantly improves diagnostic performance in the classification of breast masses: a multicenter study[J]. *Chin Med J (Engl)*, 2021, 134(4): 415-424.
- [14] Li, J., Y. Bu, S. Lu, et al. Development of a Deep Learning-Based Model for Diagnosing Breast Nodules With Ultrasound[J]. *J Ultrasound Med*, 2021, 40(3): 513-520.
- [15] Chen, C., Y. Wang, J. Niu, et al. Domain knowledge powered deep learning for breast cancer diagnosis based on contrast-enhanced ultrasound videos[J]. *IEEE Trans Med Imaging*, 2021.
- [16] Al-Antari, M. A., S. M. Han, and T. S. Kim, Evaluation of deep learning detection and classification towards computer-aided diagnosis of breast lesions in digital X-ray mammograms[J]. *Comput Methods Programs Biomed*, 2020, 196: 105584.
- [17] Schaffter, T., D. S. M. Buist, C. I. Lee, et al. Evaluation of combined artificial intelligence and radiologist assessment to interpret screening mammograms[J]. *JAMA Netw Open*, 2020, 3(3): e200265.
- [18] Xie, Y., Y. Xia, J. Zhang, et al. Knowledge-based collaborative deep learning for benign-malignant lung nodule classification on chest CT[J]. *IEEE Trans Med Imaging*, 2019, 38(4): 991-1004.
- [19] Mahase, E. AI system outperforms radiologists in first reading of breast cancer screening, study claims[J]. *BMJ*, 2020. 368: m16.
- [20] Qian, X., J. Pei, H. Zheng, et al. Prospective assessment of breast cancer risk from multimodal multiview ultrasound images via clinically applicable deep learning[J]. *Nat Biomed Eng*, 2021.
- [21] K He, X. Z., S. Ren, J. Sun, Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, 2016: 770-778.
- [22] Sanghyun Woo, J. P., Joon-Young Lee, In So Kweon, CBAM: Convolutional Block Attention Module. *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018: 3-19.
- [23] Cools-Lartigue, J. and S. Meterissian. Accuracy of axillary ultrasound in the diagnosis of nodal metastasis in invasive breast cancer: a review[J]. *World J Surg*, 2012, 36(1): 46-54.
- [24] Rao, A. A., J. Feneis, C. Lalonde, et al. A pictorial review of changes in the BI-RADS fifth edition[J]. *Radiographics*, 2016, 36(3): 623-639.
- [25] Masud, M., A. E. Eldin Rashed, and M. S. Hossain. Convolutional neural network-based models for diagnosis of breast cancer[J]. *Neural Comput Appl*, 2020: 1-12.
- [26] Nam, J. G., M. Kim, J. Park, et al. Development and validation of a deep learning algorithm detecting 10 common abnormalities on chest radiographs[J]. *Eur Respir J*, 2021, 57(5).
- [27] Xie, Y., J. Zhang, and Y. Xia, Semi-supervised adversarial model for benign-malignant lung nodule classification on chest CT[J]. *Med Image Anal*, 2019, 57: 237-248.
- [28] Jin, C., W. Chen, Y. Cao, et al. Development and evaluation of an artificial intelligence system for COVID-19 diagnosis[J]. *Nat Commun*, 2020, 11(1): 5088.
- [29] Koh, J., Y. Yoon, S. Kim, et al. Deep learning for the detection of breast cancers on chest computed tomography[J]. *Clin Breast Cancer*, 2021.
- [30] Castelvetti, D. Can we open the black box of AI? [J]. *Nature*, 2016, 538(7623): 20-23.