

## ·综述·

## 机器学习算法在辅助超声诊断附件肿块良恶性中的应用研究进展

何 新, 陈 慧, 冯炜炜

(上海交通大学医学院附属瑞金医院妇产科, 上海 200025)

**[摘要]** 卵巢癌是主要的附件恶性肿瘤, 是全球女性癌症相关死亡的第二大常见原因, 约 75% 的患者被诊断时已属晚期, 5 年生存率低于 45%。因此, 对附件肿块进行准确的非侵入性良恶性鉴别, 对于患者的预后及生存质量至关重要。近年来, 人工智能领域进展迅速, 机器学习作为人工智能领域的一个分支, 具有从大量复杂数据中进行高效学习的能力, 逻辑回归、人工神经网络、支持向量机、深度卷积神经网络等算法已被应用于辅助超声鉴别附件肿块良恶性诊断中, 并具有良好的诊断效能。本文将对机器学习算法在辅助超声鉴别附件肿块良恶性中应用价值的研究进展进行综述。

**关键词:** 附件肿块; 卵巢恶性肿瘤; 超声评分体系; 诊断模型; 机器学习

中图分类号: R541.7 文献标志码: C 文章编号: 1671-2870(2022)04-0541-06

DOI:10.16150/j.1671-2870.2022.04.022

**Research progress on the application of machine learning in assisted ultrasound diagnosis of adnexal masses HE Xin, CHEN Hui, FENG Weiwei. Department of Obstetrics and Gynecology, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, China**

**[Abstract]** Ovarian cancer is the second most common cause of gynecologic cancer death in women around the world. Around 75% of patients present with stage III/IV disease at diagnosis, with five-year survival rates below 45%. Ovarian cancer is the main adnexal malignant mass. Thus, accurate non-invasive risk stratification of adnexal masses is essential for optimal management and outcomes. In recent years, the field of artificial intelligence is developing rapidly. As a branch of artificial intelligence, machine learning could learn efficiently from complex and large amounts of data, which has infinite potential to differentiate benign and malignant adnexal masses. Logistic regression(LR), artificial neural network (ANN), support vector machine (SVM), deep learning convolution neural networks (DCNN) have been widely applied in this field, and achieved good diagnostic performance. This paper will review the history and progress of machine learning in ultrasound diagnosis of benign and malignant adnexal masses.

**Key words:** Adnexal mass; Ovarian malignant tumor; Ultrasound scoring system; Diagnostic model; Machine learning

卵巢癌是主要的附件恶性肿瘤, 居全球女性癌症相关死亡原因的第二位, 约 70% 的卵巢癌患者诊断时已属晚期<sup>[1-2]</sup>, 5 年生存率低于 45%。因此, 早期准确诊断对于卵巢癌患者治疗方案的选择以及预后改善有着重要作用。目前, 经阴道超声检查是诊断附件包块的标准一线影像学检查方法, 具有无创、经济、无辐射等主要优点<sup>[3]</sup>。超声专家评估是临床实践中诊断附件包块的最佳方法<sup>[4]</sup>。超声检查结果有较强的主观依赖性<sup>[5]</sup>, 而建立超声诊断模型则可以对附件包块进行客观分类。国际卵巢肿瘤分析小组 (The International Ovarian Tumor Analysis, IOTA) 在早期出台了关于术语和定义方面的指导建议<sup>[6]</sup>后, 在 2005 年基于逻辑回归开发了逻辑回归(logistic regression, LR) 模型 LR1 和 LR2<sup>[7]</sup>, 2008 年又

组织设计了观察超声特征的简单法则模型, 定义明确, 诊断客观、简便, 即使是初级超声学习者也可一一对照各项指标对卵巢肿块进行诊断<sup>[8]</sup>。此后, IOTA 又提出了 ADNEX (Assessment of Different NEoplasias in the adneXa) 模型, 该模型可对卵巢肿块的良恶性作出准确性较高的多分类诊断 (包括良性、交界性、早期恶性、晚期恶性和转移癌)<sup>[9]</sup>。2008 年 Amor 等<sup>[10]</sup>借鉴乳腺肿瘤影像报告与数据系统, 建立了妇科影像报告与数据系统, 作为实现附件肿块结构化报告的一种方法。2020 年, 美国放射协会提出的卵巢-附件超声报告和数据系统的风险分层与管理 (ovarian-adnexal reporting and data system, O-RADS) 共识指南, 采用规范化的超声术语, 使附件肿块的诊断报告更加规范、客观<sup>[11]</sup>。

机器学习是人工智能的一个分支, 是利用统计学(寻找数据关系)和计算机科学(有效的计算机算法)处理庞大数据集及寻找数据间内在数据关系的一种方法<sup>[12]</sup>。机器学习的过程大致分为数据收集、数据及标签预处理、模型训练及优

基金项目: 国家自然科学基金面上项目(82172601);  
上海市科委医学创新项目(20Y11914000)  
通信作者: 冯炜炜 E-mail: fww12066@rjh.com.cn

化、模型评价<sup>[13]</sup>。在构建机器学习辅助超声诊断模型时,机器学习算法可将超声图片转换为特征(像素)和标签(良性/恶性),然后使用从观察中学习的算法,计算机确定如何执行从特征到标签的映射,以创建一个模型。应用该模型进行诊断时,计算机将对新输入图像中的信息进行概括,从而自动标记肿块的良恶性。这种类似的通过计算机建立模型进行肿块良恶性诊断的方法,已经在乳腺、甲状腺、妇产科、心胸科等多个学科疾病诊断中得到了较为广泛的应用,甚至达到了与超声专家一致的诊断效能水平<sup>[14-19]</sup>。

目前,将机器学习应用于辅助超声诊断附件肿块良恶性的研究中,主要采取的算法包括 LR、人工神经网络(Artificial Neural Network, ANN)、支持向量机(support vector machines, SVM)、深度卷积神经网络(Deep learning Convolutional Neural Networks, DCNN)。近年来,机器学习算法的更新发展日新月异,其优异的诊断效能与一旦大规模应用后可极大节省医疗人力劳动成本等特点值得我们重点关注。本文将对以上常用机器学习算法模型的建立,及其在辅助超声诊断附件肿块良恶性中价值的研究进展进行综述。纳入研究汇总及算法优缺点比较见表 1、2,机器算法模式图见图 1。

#### LR 模型辅助常规超声诊断附件肿块良恶性的价值

LR 算法最早由比利时数学家维尔赫斯特在 1845 年提出,是一种用于解决二分类(0 或 1)问题的统计函数方法,用于估计某种事物的可能性<sup>[20]</sup>。LR 算法通过对数线性函数的形式将待分类数据映射到分类概率,并通过最大似然估计的方法给出结果,其假设函数形式为,其中模型的输入,为模型中调整的参数<sup>[21]</sup>。目前,LR 算法模型已在统计学、化学、计量经济学以及人工智能机器学习等多个领域中得到了普遍应用。



图 1 机器学习机器代表算法模式图

#### 一、LR1 与 LR2 模型比较

2005 年,IOTA 收集了来自 9 个中心 1 066 例患者的资料,提出 LR1 和 LR2 模型,用以评估附件肿块的恶性概率。LR1 模型通过单因素分析,从数据库中筛选出了 12 个可以用于判断附件肿块良恶性的独立变量,包括①卵巢癌个人

史(是 1,否 0);②当前激素治疗(是 1,否 0);③患者年龄(年);④病变最大直径(mm);⑤检查期间是否存在疼痛(是 1,否 0);⑥是否存在腹水(是 1,否 0);⑦在实质性乳头内是否存在血流(是 1,否 0);⑧是否存在纯实质性肿瘤(是 1,否 0);⑨实质性成分的最大直径(以毫米表示,超过 50 mm 以 50 计);⑩不规则的内囊壁(是 1,否 0);⑪存在声影(是 1,否 0);⑫血流评分(1~4 分),并对这 12 个变量进行多元回归分析以评估附件肿块的恶性概率。研究者对回归模型 LR2 进行了进一步简化,纳入的变量简化后仅包括患者年龄、腹水、是否有带血流的乳头、实质性部分的最大径线、不规则的囊肿内壁、后方是否有声影。有研究对 LR1 和 LR2 模型进行比较分析后发现,LR1 模型在术前鉴别附件肿块良恶性的效能良好,灵敏度和特异度分别为 93%、77%,受试者操作特征曲线的曲线下面积(area under curve, AUC)为 0.936;而 LR2 模型简单、易操作,但其诊断附件肿块良恶性的灵敏度和特异度稍逊,分别为 92%、75%,AUC 为 0.916。研究证明,LR 模型鉴别附件肿块良恶性的灵敏度较高,可作为门诊筛查卵巢癌的方法,但特异度欠佳,后续需要联合其他方法进一步对肿块的良恶性进行定性诊断<sup>[7]</sup>。

#### 二、LR2 与超声医师、IOTA 简单法的比较

2016 年一项荟萃分析纳入了 47 个研究,共计 19 674 例附件包肿块患者,比较超声科专家医师评估、单纯 IOTA 简单法则评估和单纯 LR2 回归模型评估在手术前诊断附件肿块良恶性的准确率。其中,简单法则模型采用的是 5 个良性特征(单房囊肿、实质性成分最大径<7 mm、伴声影、囊壁光滑的多房囊肿,最大径<10 cm、无血流信号)及 5 个恶性特征(不规则实质性肿瘤、腹水、乳头状结构≥4 个、不规则多房实质性肿瘤,最大径≥10 cm、丰富血流信号)进行评估,将肿块三分类为良性、恶性、不确定性。结果显示,超声科专家医师评估的诊断效能最优[灵敏度为 93%(95%CI 为 92%~95%),特异度为 89%(95%CI 为 86%~92%)],LR2 模型其次[灵敏度为 93%(95%CI 为 89%~95%),特异度为 84%(95%CI 为 78%~89%)]。该结果提示,在一些基层医院,超声科医师的诊断经验有限,且无条件开展外院超声科专家医师会诊支持的情况下,可以采用 LR2 模型的超声计算机学习算法辅助超声科医师在超声检查中对附件肿块良恶性进行更好地判别<sup>[22]</sup>。

LR 模型是最早被用于附件肿块良恶性诊断的计算机模型辅助诊断系统,其优势在于计算简单,可解释性强,对于数据中的每个特征,都可以通过线性部分的系数来衡量其对分类标签的影响。但由于该模型高度依赖线性关系,在一些非线性数据(如超声图像等)分析中表现欠佳,使得应用逻辑该模型的辅助诊断效能往往落后于一些新兴算法,诸如 ANN 一类,故现在的研究中一般仅将 LR 模型作为对照组使用<sup>[23-25]</sup>。

#### ANN 模型辅助常规超声诊断

##### 附件肿块良恶性的价值

ANN 是 20 世纪 80 年代以来人工智能领域兴起的研究

热点,其是基于由心理学家 McCulloch 和数理逻辑学家 Pitts<sup>[26]</sup>提出并建立的神经元数学模型。ANN 是以大脑的生理研究成果为基础,目的是模拟大脑的某些机制和机制来实现某些功能。一个 ANN 由大量神经元作为网络节点彼此连接,神经元之间通过网络、权值控制连接强度,通过调整这些连接强度,将输入数据中包含的模式信息映射到输出层。

多个研究团队对 ANN 模型诊断附件肿块良恶性的效能与既往常用的 LR 模型进行了比较<sup>[23-25]</sup>,结果发现 ANN 模型可以更好地发现变量之间可能存在的非线性相关性或相互作用。Biagiotti 等<sup>[23]</sup>分析了 226 例附件肿块患者的术前评估结果,收集了年龄、包块直径、多房与否、乳头状突起存在与否、回声分类、血流有无、收缩期峰值速度和阻力指数这 8 个预测变量,使用了三层网络和一个反向传播算法,结果显示,相较于 LR 模型,ANN 模型具有显著更高的诊断灵敏度(96%比 84%, $P=0.04$ )及类似的特异度(97.7%比 96.6%, $P=0.48$ )。

Timmermann 等<sup>[24]</sup>的研究使用年龄、绝经状态、血清 CA 125 水平和附件肿块的超声特征作为变量,比较了 LR 模型与 ANN 模型在评估 173 个卵巢肿瘤样本恶性概率中的效能,同样发现通过 ANN 模型进行诊断的准确率(AUC 为 0.979,灵敏度为 95.9%,特异度为 93.5%)要明显高于传统的 LR 模型(AUC 为 0.956,灵敏度为 95.9%,特异度为 85.5%)。

Moszynski 等<sup>[25]</sup>使用患者附件肿块的灰阶和多普勒超声参数,联合其临床信息及血清肿瘤标志物水平,分别应用了 ANN 和 LR 模型进行分析,结果发现,ANN 的诊断效能(灵敏度为 85.7%,特异度为 93.1%)高于 LR 模型(灵敏度为 65.71%,特异度为 77.59%)。

以上研究均证明,应用 ANN 的机器学习算法较传统的 LR 模型能更有效地评估所检查的附件肿块特征。ANN 算法在相对模糊、重叠的诊断类别之中更高效、准确。ANN 算法的出现一度引起研究热潮,但其也存在以下缺点。①建立 ANN 模型一般需要大量的训练样本;②代价函数边界复杂,存在多个局部最优值;③浅层 ANN 对于特征学习的表达能力有限,而深层 ANN 的参数繁多,易导致过拟合,且有可能因为梯度消失而导致不可学习。因此,研究者们提出了新的模型算法——SVM。

### SVM 模型辅助常规超声诊断

#### 附件肿块良恶性的价值

SVM 是基于统计学习理论发展而来,本质上依然是一个线性分类器,而其核心思想在于其分类准则为最大间隔,即要将两类分开,寻找一个满足所要求的分类条件的最优分类超平面,该超平面与任何类的最近的训练数据点具有最大距离,因为通常边距越大,泛化误差越小,即所得的分类器有很好的通用性<sup>[27]</sup>。与之前的机器学习算法如 ANN 算法相比,SVM 算法不存在陷入局部最小值的风险及过拟合的问题,在处理回归问题和分类以及判别问题时能起到较好的效果。

纹理是影像学计算机辅助诊断系统中常用的特征之一,其将可肉眼观察到的图像经过计算而量化,变化的图像灰度分布特点可以用于区分图像内容<sup>[28]</sup>。SVM 算法在应用于附件肿块良恶性诊断时,多由检查者直接从超声图像中提取纹理特征(如局部二值模式、灰度共生矩阵、纹理能量和 Gabor 小波等),继而结合 SVM 分类器以完成肿瘤良恶性鉴别。Acharya 等<sup>[29]</sup>使用局部二值模式等纹理特征算子,从卵巢的 3D 超声图像中提取特征,并结合 SVM 来鉴别卵巢肿瘤的良恶性。他们在 SVM 分类器中分别导入从 1 000 个良性卵巢肿瘤和 1 000 个恶性卵巢肿瘤超声图像中提取的纹理能量和局部二值模式的纹理特征,在最终的结果验证中获得了高达 99.9% 的准确率、100% 的灵敏度和 99.8% 的特异度。但该研究纳入的几千张图像均是从 20 例卵巢肿瘤患者(10 例良性肿瘤、10 例恶性卵巢肿瘤)的检查图像中转换得到的,在外推过程中可能存在比较大的样本偏倚。

SVM 模型也存在以下缺点。第一,常规 SVM 算法的分类结果为二分类,即只能提供最后判断结果是(1)或否(0),而不能直接提供概率估计。第二,只有当核函数与数据的分布较为吻合时才能得到好的效果。Khazendar 等<sup>[28]</sup>的研究基于来自 177 例比利时鲁汶大学医院妇产科的病例(共 187 个图像),建立了 SVM 决策层融合策略,结果发现其中 18.3% 的图像不能被该系统分类。若仅用特征进行评估的平均准确率仅为 77%<sup>[30]</sup>。说明该算法还有待进一步改进,其辅助诊断效能还有待进一步提高。

### DCNN 模型辅助常规超声诊断

#### 附件肿块良恶性的价值

DCNN 是一类包含卷积计算且具有深度结构的神经网络,是深度学习(Deep Learning, DL)的代表算法之一。对 DCNN 的研究始于 20 世纪 80 年代至 90 年代。得益于 DL 理论的提出和数值计算设备的改进,DCNN 得到了快速发展,并被大量应用于计算机视觉、自然语言处理等领域<sup>[31]</sup>。2015 年后,大量基于 DCNN 算法的解决方案开始出现,并已成功用于一些医学影像学检查的辅助诊断中。DL 算法在医学影像辅助诊断系统中的应用具有以下优点。首先,作为一种数据驱动的自动特征学习算法,其可以直接从训练数据中提取特征,从而不像其他机器学习算法那样需要人工进行病例数据特征的设计和计算,大大减少了特征提取的工作量和人工偏倚的影响<sup>[32]</sup>。其次,DL 算法可以通过神经网络的深层结构来表示特征的交互。并且,在端到端的优化过程中,其可同时实现特征提取、特征选择和分类 3 个核心步骤。基于 DCNN 算法在医学影像学领域的成功应用经验,利用 DCNN 模型对卵巢肿瘤超声图像进行分类,从而使早期判断卵巢肿瘤良恶性成为可能。

天津市中心妇产科医院 Zhang 等<sup>[33]</sup>在 2019 年提出了一种用于辅助彩超诊断卵巢肿瘤良恶性的计算机辅助诊断系统。该团队创新性地联合分析了基于 DL 网络特征和基于纹理描述符特征所融合的影像特点。深度神经网络从医学超

声图像中提取的特征可反映病变区域，而纹理特征可以描述病变区域的边缘、方向和强度分布。结果表明，结合2种特征可有效区分病变区域与其他区域，恶性卵巢肿瘤与良性卵巢囊肿病变区域。最终的结果验证显示，在该计算机辅助诊断系统的辅助下，彩超诊断卵巢肿瘤良恶性的准确率达到了99.15%，灵敏度达99.73%，特异度达95.85%，AUC为0.997。

DL算法需要大量的高质量标注数据，且不说大量数据来源困难，标注数据亦是一项工程量巨大的任务，这时候采用迁移学习是一个较好的解决方案。迁移学习是指将原领域中预先训练好的部分网络，包括其网络结构和连接参数，重新利用，使之转化为用于目标领域的深度神经网络的一部分，尤其是以图像领域为代表，很多情况下会选择预训练模型来对预测模型进行初始化。Christiansen等<sup>[34]</sup>通过迁移学习预训练模型VGG16<sup>[35]</sup>、ResNet50<sup>[36]</sup>和MobileNet<sup>[37]</sup>，整合建立了Ovry-Dx1(测试集中的所有病例被分类为良性或恶性)及Ovry-Dx2(测试集中的病例被分类为良性\恶性\不确定性)2个DNN模型，来评估卵巢肿瘤的性质，并将该DNN模型的诊断效能与经验丰富的超声专家的诊断结果进行比较。结果显示，Ovry-Dx1模型的诊断灵敏度为96.0%，特异度

86.7%，与经验丰富的临床专家相似(灵敏度为96.0%，特异度88.0%， $P=1.0$ )。在不确定病例比例为12.7%时，Ovry-Dx2模型的诊断灵敏度为97.1%，特异度为93.7%。在不确定病例中，采用Ovry-Dx2模型辅助专家进行评估，总的灵敏度(96.0%)和特异度(89.3%)得到了明显提升，可达到与专家同等水平( $P=1.0$ )。

同济医院Gao等<sup>[38]</sup>纳入分析了2003年9月至2019年5月期间，十家医院共107 624名女性(包括4 254例卵巢癌患者和103 370名良性附件肿块患者)的卵巢超声病例资料，共提取导入了总计592 275张超声图像。结果发现，DCNN模型诊断卵巢癌的AUC高达0.911，该模型在内部及外部验证数据集中诊断卵巢肿瘤良恶性的准确率(88.8%、86.9%)对比35名经验丰富的超声医师的平均诊断准确率更高(85.7%、81.1%)。采用DCNN辅助后超声诊断卵巢肿瘤良恶性的准确率和灵敏度均较常规超声检查有所提高(87.6%比78.3%， $P<0.0001$ ；82.7%比70.4%， $P<0.0001$ )，6名初级超声科医生的DCNN辅助评估的平均准确度达到0.876( $P<0.05$ )。

笔者所在的研究团队曾对本中心2019年1月至2019年11月期间接受超声检查及手术治疗的卵巢肿瘤患者进行回

表1 文中纳入机器学习在超声诊断附件包块良恶性中的研究汇总

	研究	年份	AUC	灵敏度	特异度	图像数量(恶性/良性)	患者例数(恶性/良性)	模型
LR	Timmerman <sup>[7]</sup>	2005	0.936	93%	77.0%		1 066(266/800)	LR1
			0.916	92%	75.0%			LR2
ANN	Biagiotti <sup>[23]</sup>	1999		96%	97.7%		226(51/175)	
	Timmermann <sup>[24]</sup>	1999	0.979	95.9%	93.5%			173(49/124)
	Moszynski <sup>[25]</sup>	2006	0.968	85.7%	93.1%			686(255/431)
SVM	Acharya <sup>[29]</sup>	2012		100.0%	99.8%	2 000(1 000/1 000)	20(10/10)	SVM, RBF
				99.6%	100%			SVM, linear
CNN	Khazendar <sup>[28]</sup>	2019		80%	77.0%	187(75/112)	177	
				99.73%	95.85%			
	Zhang <sup>[33]</sup>	2019	0.997			428(357/71)	758(309/449)	Ovry-Dx1
			0.950	96.0%	86.7%			Ovry-Dx2
	Christiansen <sup>[34]</sup>	2021		0.958	97.1%	3 077	107 624(4 254/103 370)	
					93.7%			
Gao <sup>[38]</sup>		2022	0.911	83.1%	86.8%	592 275(39 258 /553 017)	422(118/304)	DL <sub>feature</sub>
			0.930	92%	85.0%			DL <sub>decision</sub>
			0.900	92%	80.0%			

表2 文中所述机器学习算法优缺点比较

算法	优点	缺点
逻辑回归算法	(1)便于理解和实现，可以观测样本的概率分数。 (2)训练速度快	(1)容易欠拟合。 (2)在一些非线性的数据上表现欠佳
人工神经网络	(1)分类的准确度高，学习能力强。 (2)有较强的鲁棒性和容错能力。	(1)需要大量的训练样本及参数。 (2)浅层神经网络对于特征学习的表达能力有限，深层神经网络的参数繁多，易导致过拟合，甚至可能因为梯度消失而导致不可学习。
支持向量机	(1)SVM是一种小样本学习方法。 (2)算法简单，而且具有较好的鲁棒性。	(1)常规SVM的分类结果是二分类的，用SVM解决多分类问题存在困难，且不能直接提供概率估计。 (2)SVM算法只有当核函数与数据的分布较为吻合时才能得到好的效果
深度卷积神经网络	(1)具有自学习功能，可以直接从训练数据中提取特征。 (2)可以通过神经网络的深层结构来表示特征之间的关系。 (3)能够获取大量数据中包含的信息，同时实现特征提取、特征选择和分类3个核心步骤，并构建模型。	(1)具有黑盒性，目前研究无法确定DCNN网络内部是如何运行的。 (2)网络模型复杂程度越高，其对计算设备的硬件要求就越高，对于数据质量的要求也更高

顾性研究<sup>[39]</sup>,采用 2 种融合策略,开发了 2 种基于残差网络(ResNet)的 DL 模型,其一为特征融合模型(DLfeature),其二为决策融合模型(DLdecision),分别进行卵巢肿瘤良恶性评估,并将 DL 模型预测结果与 O-RADS 风险分级、常规超声检查(专家评估)结果进行比较。结果显示,DLfeature 模型诊断卵巢肿瘤良恶性的 AUC 为 0.93(95%CI 为 0.85~0.97),与 O-RADS 风险分级(AUC 为 0.92; 95% CI 为 0.85, 0.97; P=0.88)、常规超声检查(AUC 为 0.97; 95%CI 为 0.91, 0.99; P=0.07) 及 DLdecision 模型(AUC 为 0.90; 95%CI 为 0.82, 0.96; P=0.29)相似。DLdecision 模型、DLfeature 模型、O-RADS 风险分级和常规超声诊断卵巢恶性肿瘤的灵敏度分别为 92%、92%、92% 和 96%,特异度分别为 80%、85%、89% 和 87%。这说明了使用多模态超声图像开发的 DL 算法可以较好地区分恶性与良性卵巢肿瘤,其诊断性能可与专家主观评估相媲美。

但 DCNN 算法同样具有一定的缺点。首先,DCNN 算法具有黑盒性,目前研究无法确定 DCNN 网络内部是如何运行的,尽管 DCNN 算法可以在许多任务上实现卓越的性能,但缺乏可分解性,无法分解为直观的组件,因此难以解释。第二,DCNN 算法的使用成本较高,网络模型复杂程度越高,其对计算设备的硬件要求就越高,对于数据质量的要求也更高。所以针对具体的数据集采用最适用的模型和算法,从而达到在成本和诊断准确率上的平衡,也是研究者们需要考虑的。

## 小 结

机器学习辅助超声诊断提供了一种基于大数据的客观鉴别诊断附件肿块良恶性的途径。希望在不久的将来,随着计算机技术的日新月异,机器学习辅助超声诊断可以为附件包块早期诊断、精准治疗提供参考。

## 【参考文献】

- [1] Webb PM, Jordan SJ. Epidemiology of epithelial ovarian cancer[J]. Best Pract Res Clin Obstet Gynaecol,2017,41:3-14.
- [2] Lheureux S, Braunstein M, Oza AM. Epithelial ovarian cancer: Evolution of management in the era of precision medicine[J]. CA Cancer J Clin,2019,69(4):280-304.
- [3] Fischerova D. Ultrasound scanning of the pelvis and abdomen for staging of gynecological tumors: a review [J]. Ultrasound Obstet Gynecol,2011,38(3):246-266.
- [4] Tavoraitè I, Kronlachner L, Opolskienė G, et al. Ultrasound Assessment of Adnexal Pathology: Standardized Methods and Different Levels of Experience[J]. Medicina (Kaunas),2021,57(7):708.
- [5] Timmerman D, Planchamp F, Bourne T, et al. ESGO/ISUOG/IOTA/ESGE Consensus Statement on pre-operative diagnosis of ovarian tumors[J]. Int J Gynecol Cancer, 2021,31(7):961-982.
- [6] Timmerman D, Valentin L, Bourne TH, et al. Terms, definitions and measurements to describe the sonographic features of adnexal tumors: a consensus opinion from the International Ovarian Tumor Analysis (IOTA) Group[J]. Ultrasound Obstet Gynecol,2000,16(5):500-505.
- [7] Timmerman D, Testa AC, Bourne T, et al. Logistic regression model to distinguish between the benign and malignant adnexal mass before surgery: a multicenter study by the International Ovarian Tumor Analysis Group [J]. J Clin Oncol,2005,23(34):8794-8801.
- [8] Timmerman D, Testa AC, Bourne T, et al. Simple ultrasound-based rules for the diagnosis of ovarian cancer[J]. Ultrasound Obstet Gynecol,2008,31(6):681-690.
- [9] van Calster B, van Hoorn K, Valentin L, et al. Evaluating the risk of ovarian cancer before surgery using the ADNEX model to differentiate between benign, borderline, early and advanced stage invasive, and secondary metastatic tumours: prospective multicentre diagnostic study[J]. BMJ,2014,349:g5920.
- [10] Amor F, Vaccaro H, Alcázar JL, et al. Gynecologic imaging reporting and data system: a new proposal for classifying adnexal masses on the basis of sonographic findings [J]. J Ultrasound Med,2009,28(3):285-291.
- [11] Andreotti RF, Timmerman D, Strachowski LM, et al. O-RADS US Risk Stratification and Management System: A Consensus Guideline from the ACR Ovarian-Adnexal Reporting and Data System Committee[J]. Radiology,2020,294(1):168-185.
- [12] Awad M, Khanna R. Machine Learning [M]// Efficient Learning Machines. CA:Apress,Berkeley,2015:1-18.
- [13] Rajkomar A, Dean J, Kohane I. Machine Learning in Medicine[J]. N Engl J Med,2019,380(14):1347-1358.
- [14] Brattain LJ, Telfer BA, Dhyani M, et al. Machine learning for medical ultrasound: status, methods, and future opportunities[J]. Abdom Radiol(NY),2018,43(4):786-799.
- [15] Moustafa AF, Cary TW, Sultan LR, et al. Color Doppler Ultrasound Improves Machine Learning Diagnosis of Breast Cancer[J]. Diagnostics (Basel),2020,10(9):631.
- [16] Zhao CK, Ren TT, Yin YF, et al. A Comparative Analysis of Two Machine Learning-Based Diagnostic Patterns with Thyroid Imaging Reporting and Data System for Thyroid Nodules: Diagnostic Performance and Unnecessary Biopsy Rate[J]. Thyroid,2021,31(3):470-481.
- [17] Sone K, Toyohara Y, Taguchi A, et al. Application of artificial intelligence in gynecologic malignancies: A review [J]. J Obstet Gynaecol Res,2021,47(8):2577-2585.
- [18] Xie HN, Wang N, He M, et al. Using deep-learning algo-

- rithms to classify fetal brain ultrasound images as normal or abnormal[J]. *Ultrasound Obstet Gynecol*,2020,56(4):579-587.
- [19] Al'Aref SJ, Anchouche K, Singh G, et al. Clinical applications of machine learning in cardiovascular disease and its relevance to cardiac imaging[J]. *Eur Heart J*,2019,40(24):1975-1986.
- [20] Verhulst PF. Recherches mathématiques sur la loi d'acroissement de la population[J]. Nouveaux mémoires de l'académie royale des sciences et belles-lettres de Bruxelles,1845,18:1-38.
- [21] Namburi S. Logistic regression with conjugate gradient descent for document classification[D]. Kansas State University,2016.
- [22] Meys EM, Kajser J, Kruitwagen RF, et al. Subjective assessment versus ultrasound models to diagnose ovarian cancer: A systematic review and meta-analysis[J]. *Eur J Cancer*,2016,58:17-29.
- [23] Biagiotti R, Desii C, Vanzi E, et al. Predicting ovarian malignancy: application of artificial neural networks to transvaginal and color Doppler flow US[J]. *Radiology*,1999,210(2):399-403.
- [24] Timmerman D, Verrelst H, Bourne TH, et al. Artificial neural network models for the preoperative discrimination between malignant and benign adnexal masses[J]. *Ultrasound Obstet Gynecol*,1999,13(1):17-25.
- [25] Moszynski R, Szpurek D, Smolen A, et al. Comparison of diagnostic usefulness of predictive models in preliminary differentiation of adnexal masses[J]. *Int J Gynecol Cancer*,2006,16(1):45-51.
- [26] McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. 1943[J]. *Bull Math Biol*, 1990,52(1-2):99-115.
- [27] Moguerza JM, Muoz Alberto. Support Vector Machines with Applications 1[J]. *Statistical Science*,2006,21(3):358-362.
- [28] Khazendar S, Sayasneh A, Al-Assam H, et al. Automated characterisation of ultrasound images of ovarian tumours: the diagnostic accuracy of a support vector machine and image processing with a local binary pattern operator[J]. *Facts Views Vis Obgyn*,2015,7(1):7-15.
- [29] Acharya UR, Sree SV, Krishnan MM, et al. Ovarian tumor characterization using 3D ultrasound[J]. *Technol Cancer Res Treat*,2012,11(6):543-552.
- [30] Khazendar S, Al-assam H, Du H, et al. Automated classification of static ultrasound images of ovarian tumors based on decision level fusion[C]. proceedings of the 2014 6th Computer Science and Electronic Engineering Conference (CEEC),2014:25-26.
- [31] Khan A, Sohail A, Zahoor U, et al. A survey of the recent architectures of deep convolutional neural networks [J]. *Artificial intelligence review*,2020,53(8):5455-5516.
- [32] Cai L, Gao J, Zhao D. A review of the application of deep learning in medical image classification and segmentation[J]. *Ann Transl Med*,2020,8(11):713.
- [33] Zhang L, Huang J, Liu L. Improved Deep Learning Network Based in combination with Cost-sensitive Learning for Early Detection of Ovarian Cancer in Color Ultrasound Detecting System[J]. *J Med Syst*,2019,43(8):251.
- [34] Christiansen F, Epstein EL, Smedberg E, et al. Ultrasound image analysis using deep neural networks for discriminating between benign and malignant ovarian tumors: comparison with expert subjective assessment[J]. *Ultrasound Obstet Gynecol*,2021,57(1):155-163.
- [35] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[C]. in International Conference on Learning Representations,2015.
- [36] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR),2016.
- [37] Howard AG, Zhu M, Chen B, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications[J]. arXiv:1704.04861,2017.
- [38] Gao Y, Zeng S, Xu X, et al. Deep learning-enabled pelvic ultrasound images for accurate diagnosis of ovarian cancer in China: a retrospective, multicentre, diagnostic study[J]. *Lancet Digit Health*,2022,4(3):e179-e187.
- [39] Chen H, Yang BW, Qian L, et al. Deep Learning Prediction of Ovarian Malignancy at US Compared with ORADS and Expert Assessment[J]. *Radiology*,2022,304(1):106-113.

(收稿日期:2022-02-20)

(本文编辑:褚敬申)