

基于深度卷积神经网络的甲状腺超声图像 良恶性结节识别方法研究

姚文君^{1,2},殷超然³,朱宏庆¹,江健敏⁴,庞小溪⁵,孙怡宁²

摘要 目的 探讨深度卷积神经网络在甲状腺结节超声图像的自动检测和良恶性分类中应用价值。**方法** 回顾性选取 1 012 幅甲状腺结节的超声图像并对其进行标记,构建 YOLOv5 网络模型,精准定位甲状腺结节所在位置并自动裁减结节所在区域,同时构建 GoogLeNet 网络模型对裁减后结节的图像进行良恶性分类。**结果** 在所采集的数据集中,目标检测网络对甲状腺结节位置检测的平均精确度均值为 96.2%;分类网络对良恶性结节分类的敏感度为 0.885,特异度为 0.822,准确度为 0.866,AUC 值为 0.92,显著高于 Alex-Net 模型(AUC = 0.81)、VGG 模型(AUC = 0.86)和 MobileNet 模型(AUC = 0.76)。**结论** 深度卷积神经网络模型对超声图像中的甲状腺良恶性结节具有较高的定位和识别能力,有助于提高影像自动诊断的准确性。

关键词 甲状腺结节;超声图像;深度卷积神经网络;YOLOv5 网络

中图分类号 R 816.6;R 320.1140

文献标志码 A **文章编号** 1000-1492(2023)05-0854-05
doi:10.19405/j.cnki.issn1000-1492.2023.05.025

近几十年来,甲状腺癌的发病率在全世界范围内迅速上升^[1],由于早期甲状腺癌治愈的可能性极大,因此,对甲状腺良恶性结节的早期鉴别至关重要^[2]。超声虽然被广泛应用于甲状腺结节的诊断和随访,但视觉评估的图像耗时、主观性强,迫切需要一种准确、高效、客观的辅助诊断方法。目前计算机辅助诊断已有一定的研究基础^[3],而卷积神经网络

的兴起使其表现出了更多的实用价值^[4],如 GoogLeNet 常用于医学图像的分类^[5]。然而,超声图像分辨率低、对比度低,识别难度较大,且目前大部分利用手动分割,故诊断的精度和自动化水平有待提高。基于此,该研究利用 YOLOv5 神经网络的深度学习方法,构建甲状腺结节区域卷积神经网络检测模型,自动获取结节区域图像并利用 GoogLeNet 图像分类网络对结节进行良恶性分类,从而实现影像全过程的自动化诊断,有效缩短阅片时间,规避主观因素对诊断造成的影响,提高甲状腺结节的识别和诊断水平,避免过度治疗带来的损伤。

1 材料与方法

1.1 病例资料 回顾性选取安徽省第二人民医院及安徽医科大学第二附属医院 2019 年 6 月—2022 年 6 月的甲状腺超声图像,共 1 012 例,其中男性 352 例,女性 660 例。纳入标准:超声图像清晰。排除标准:超声资料不完整;同时患有其他肿瘤疾病;桥本氏甲状腺炎;甲状腺弥漫性钙化性结节;伪影较多影响图像质量。本研究经安徽医科大学第二附属医院伦理委员会通过,免除知情同意书(批准号:YX2021-001)。超声仪器为 GE S8、GE S9 及迈瑞 R7T,采用 7~12 MHz 高频线阵探头。

1.2 甲状腺结节的分类与标注 由 2 名高年资影像医师共同对结节范围进行定位标注,并进行良恶性分类,再由 1 名副主任医师对上述结果进行判别修正。良性分类标准为:TI-RADS 2 级或有 2 次以上随访无变化的 3 级结节、或确诊为良性的结节。恶性分类标准为:穿刺活检或手术后病理为恶性或 TI-RADS 4b 以上的结节。最终建立甲状腺结节数据集,共 1 012 个结节,其中 703 个标注为良性,309 个被标注为恶性。本研究将数据集按照 8:2 的比例划分为训练集与测试集,其中数据集的分布情况见表 1。

1.3 数据预处理

1.3.1 数据筛选 数据集中含有 1 012 幅原始甲状腺结节超声图像,由于不同制造商生产的各种

2022-11-29 接收

基金项目:安徽省自然科学基金(编号:2008085QH406);安徽医科大学第二附属医院临床研究培育计划(编号:2020LCYB05);安徽省转化医学研究院科研基金(编号:2021zhyx-C45)

作者单位:安徽医科大学第二附属医院¹放射科、⁵核医学科,合肥 230601

²中国科学院合肥物质研究院智能机械研究所,合肥 230031

³安徽省第二人民医院超声科,合肥 230011

⁴安徽大学电子信息工程学院,合肥 230601

作者简介:姚文君,女,副主任医师;

孙怡宁,男,研究员,博士生导师,责任作者;E-mail: ynsun@iim.ac.cn

表1 数据集分布表(个)

数据集	良性结节	恶性结节	合计
训练集	564	247	811
测试集	139	62	201
合计	703	309	1 012

型号的超声波设备生成超声图像格式不尽相同,首先将所有甲状腺超声图像转换为 JPG 格式。为了确保数据集中原始图像具有较高质量,对所有甲状腺图像进行筛选,去除含有严重伪影(如运动伪影、折射伪影等)或图像分辨率显著过低的低质量原始图像。

1.3.2 提取感兴趣区域(region of interest, ROI)

原始超声图像包含噪声信息(如超声设备的参数、原始图像的测量标尺等)通常分布在原始图像的外围部分,本研究将原始图像外围包含噪声信息的部分手动去除,并对图像进行对比度增强,以提升图像质量,突出甲状腺结节特征,帮助网络更好检测结节位置和判断结节良恶性。

1.4 模型建立

1.4.1 甲状腺结节区域检测模型 首先采用 YOLOv5 目标检测网络,构建适合甲状腺影响的甲

状腺结节区域检测模型。YOLOv5 是单阶段目标检测方法,其在模型训练以及推理时间上具有良好的表现,并且已经成为了最为广泛应用的目标检测网络之一。YOLOv5 网络包含输入端、骨干网络、颈部网络和输出端四部分,结构见图 1。

1.4.2 甲状腺结节良恶性分类模型 GoogLeNet 分类网络提出了新的 Inception 深度学习结构,可以有效解决由于深度学习网络加深带来的梯度消失和梯度爆炸等问题,Inception 结构的提出还使网络更高效的利用计算资源。而 GoogLeNet 网络中含有 9 个 Inception 结构,通过 Inception 结构的重复提取,能够在一定计算资源下提取到图像中更深层的特征。在 GoogLeNet 结构中,还含有两个辅助分类器用于向前传导梯度,避免梯度的消失。GoogLeNet 网络流程图见图 2。

1.4.3 方法流程 首先将经过预处理的甲状腺结节图像输入 YOLOv5 检测网络,输出甲状腺结节的位置,然后自动从原图像中裁剪出甲状腺结节区域,并对该区域图像进行去噪及调整大小,得到甲状腺结节图像,之后将图像送入 GoogLeNet 分类网络得到结节的分类结果。方法的总体流程见图 3。

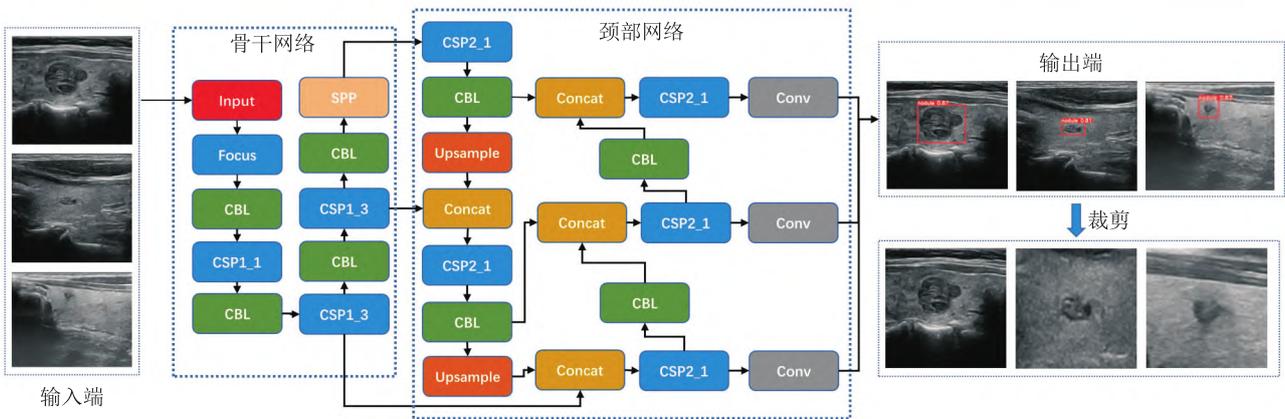


图1 YOLOv5 甲状腺结节检测网络流程图

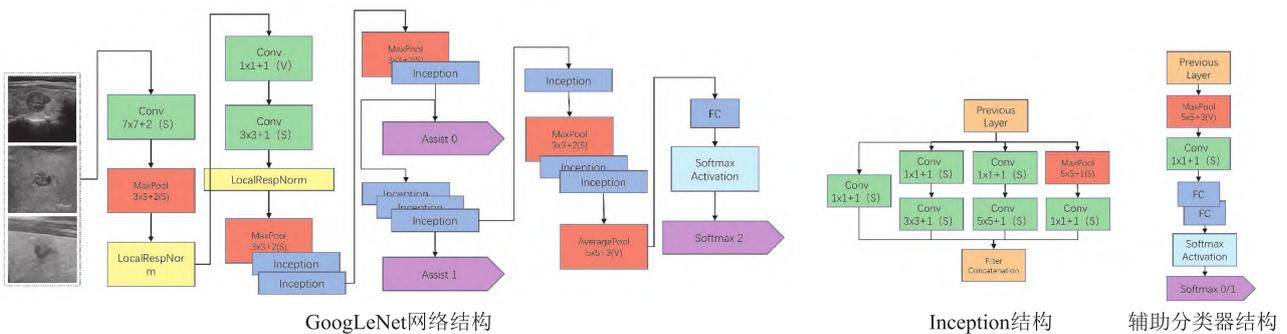


图2 GoogLeNet 甲状腺结节图像分类网络流程图

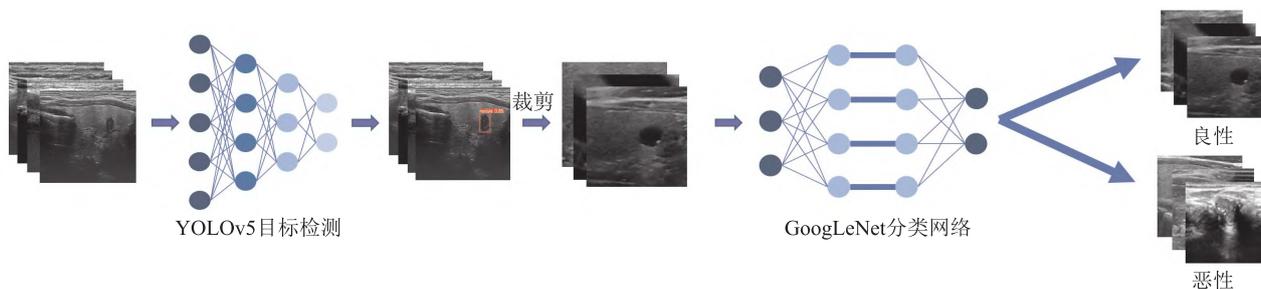


图3 方法总体流程图

1.5 评价指标 在本研究中使用了目标检测网络和分类网络。对于目标检测网络来说,常用的性能评价指标为平均精确度均值 (mean average precision, mAP), 即 mAP 的值越大其对应的目标检测模型的定位与识别性能越高。平均精确度均值 mAP 的定义公式如式(1)所示:

$$(1) mAP = \frac{\sum_{i=1}^N \int_0^1 P(R) dR}{N}$$

其中, P(R) 为检测结果的准确率与召回率曲线, N 为检测种类数。

对于分类网络来说,本研究采用敏感度 (sensitivity, SENS)、特异度 (specificity, SPEC)、准确度 (accuracy, ACC)、kappa 系数、正确率 (precision, PREC) 和曲线下面积 (area under curve, AUC) 值衡量分类网络性能。其中敏感度表示所有正例中被分类正确的比例, 特异性表示所有负例中被分类正确的比例, 准确度表示所有被分为正例中实际也为正例的比例, 正确率表示所有分对的样本占总体样本的比例。计算公式如式(2) ~ (6) 所示:

$$(2) SENS = \frac{TP}{TP + FN}$$

$$(3) SPEC = \frac{TN}{TN + FP}$$

$$(4) ACC = \frac{TP + TN}{TP + FP + FN + TN}$$

$$(5) PREC = \frac{TP}{TP + FP}$$

$$(6) KAPPA = \frac{ACC - Pe}{1 - Pe}$$

其中, TP 和 TN 分别表示被分类网络正确诊断的良性和恶性结节数, FP 和 FN 分别表示被分类网络错误诊断的良性和恶性结节数。Pe 为偶然一致性误差。

此外, 通过受试者工作特征曲线 (receiver operating characteristic, ROC) 及 AUC 评估各算法的性

能。

2 结果

2.1 甲状腺结节检测结果 本研究采用了 YOLOv5 目标检测网络对甲状腺结节的位置进行检测。首先, 利用训练集对 YOLOv5 模型进行训练, 直到损失函数收敛, 然后利用训练后的模型在测试集中检测甲状腺结节位置。结果表明, YOLOv5 检测甲状腺结节 mAP 达到了 96.2%, YOLOv5 网络对甲状腺结节的定位精度较高。训练后的模型在测试集中部分检测结果示例见图 4。

2.2 甲状腺结节的图像分类结果及一致性检验

为了评估 GoogLeNet 分类方法的性能, 将其与 AlexNet、VGG 和 MobileNet 三种常见的分类方法进行对比实验。对比实验均在本研究所构建的训练集上进行训练、测试集上获得分类结果, 各方法分类结果见表 2, 结果表明, GoogLeNet 对甲状腺结节分类的正确率达到了 0.918, 高于其它几种方法。各分类网络 ROC 曲线见图 5。

表 2 甲状腺结节分类结果对比

标签	敏感度	特异度	准确度	正确率	kappa 系数	AUC 值
AlexNet	0.885	0.774	0.851	0.898	0.653	0.81
VGG	0.892	0.725	0.841	0.879	0.624	0.86
MobileNet	0.942	0.500	0.805	0.809	0.491	0.76
GoogLeNet	0.885	0.822	0.866	0.918	0.693	0.92

3 讨论

甲状腺癌是内分泌系统最常见的恶性肿瘤之一, 发病率逐年上升。2018 年甲状腺癌诊疗规范中推荐所有触诊或机会性筛查等方式发现甲状腺结节者均应行超声检查, 从而增加了影像科医师工作强度, 缩减了每位患者的平均诊断时间, 直接影响结果的准确性, 因此临床迫切需要一种高效、准确的方法辅助影像医师提高工作效率及准确性。本研究利用

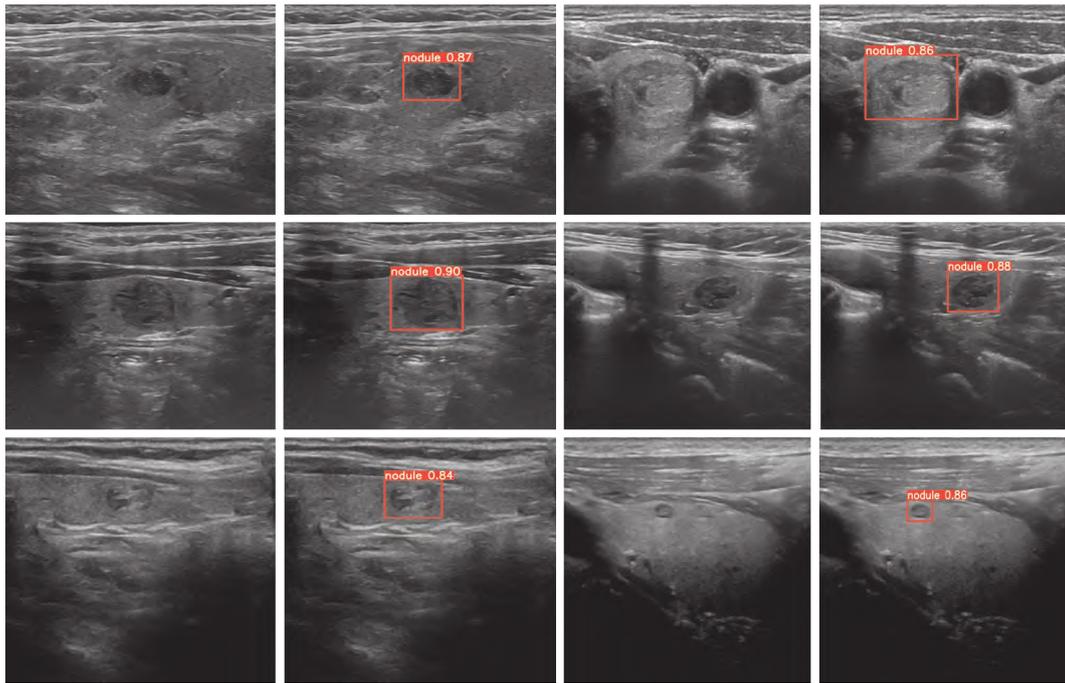


图4 YOLOv5 部分检测结果

红色矩形框为模型检测出来的甲状腺结节位置,矩形上表示检测出目标的置信度

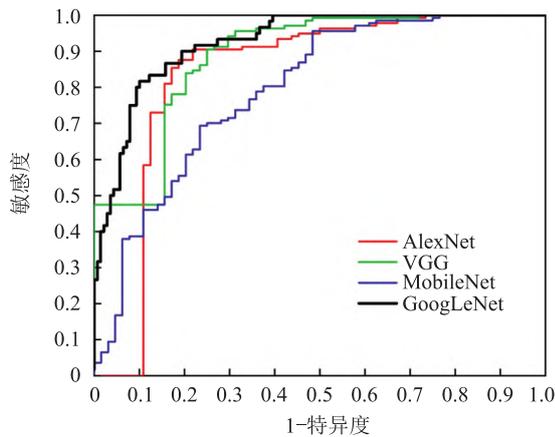


图5 各网络对甲状腺结节分类的 ROC 曲线

深度卷积神经网络对甲状腺超声图像中的结节进行自动检测和良恶性分类,结果表明,结节位置检测的 mAP 达到了 96.2%;同时对良恶性结节分类敏感度为 0.885,特异度为 0.822,准确度为 0.866,正确率为 0.918,AUC 值为 0.92,显著优于 AlexNet、VGG 及 MobileNet。

深度学习尤其是卷积神经网络已广泛应用于图像分类、目标检测等各种视觉任务中,如肺结节检测及鉴别、脑卒中及老年痴呆症的诊断、骨折与骨龄检测等,近年来在甲状腺疾病的检测、分割及鉴别等方面也取得了较好的研究进展^[6-8]。张静漪等^[9]提

出利用 Mask R-CNN 模型为基础加入级联网络,对正常及异常的甲状腺超声图像进行目标检测、分割及分类,充分显示了卷积神经网络在自动识别甲状腺超声图像方面的优势;乔婷婷等^[10]利用了 2 种深度神经网络模型 AlexNet 和深度卷积生成对抗网络对甲状腺平面显像图进行特征提取和功能判断,识别准确性较高;李潜等^[11]应用了基于卷积神经网络的辅助诊断软件 S-Detect 技术对甲状腺结节的超声图像进行了良恶性鉴别,结果表明将人工智能技术用于辅助诊断可以提高低年资医师的诊断准确率;此外,在端到端的卷积神经网络模型方面也取得了良好的诊断效能^[12]。本研究提出的基于深度卷积神经网络的甲状腺超声图像良恶性结节识别方法,首先使用了 YOLOv5 目标检测网络对结节位置进行检测,继而利用 GoogLeNet 图像分类网络对结节进行良恶性分类,此方法能在自动精确检测甲状腺结节位置的同时,对结节的良恶性进行分类,从而实现了影像全过程的自动化诊断。目前临床采用 TI-RADS 或 C-TIRADS^[13]来进行分类和评估结节恶性风险,但实际操作耗时较多,准确性也依赖操作者的经验;与之相比,基于深度卷积神经网络的模型具备很多优势,如“对图像全信息的识别和利用”、“可重复性高”、“模型随着病例数的增多会不断优化”等,因此课题组后期会继续搜集更多数据,以提高模型

效能。

本研究具有一定的局限性:人工智能系统的开发是为了辅助诊断,而不是取代影像医师,本研究只进行了模型之间的对比,并未评估影像医师在模型辅助下诊断效能提高的程度,后续研究将评估不同经验的影像医师在模型辅助下诊断效能的变化程度。

参考文献

[1] Siegel R L, Miller K D, Jemal A. Cancer statistics, 2019 [J]. *CA Cancer J Clin*, 2019, 69(1):7-34.

[2] Luo Y, Zhao Y, Chen K, et al. Clinical analysis of cervical lymph node metastasis risk factors in patients with papillary thyroid microcarcinoma [J]. *J Endocrinol Invest*, 2019, 42(2):227-36.

[3] Park H J, Kim S M, La Yun B, et al. A computer-aided diagnosis system using artificial intelligence for the diagnosis and characterization of breast masses on ultrasound: added value for the inexperienced breast radiologist [J]. *Medicine (Baltimore)*, 2019, 98(3):e14146.

[4] Gu Y, Chi J Q, Liu J Q, et al. A survey of computer-aided diagnosis of lung nodules from CT scans using deep learning [J]. *Comput Biol Med*, 2021, 137:104806.

[5] Lakshmi priya B, Jayanthi K P, Biju P, et al. GoogLeNet-based en-

semble FCNet classifier for focal liver lesion diagnosis [J]. *IEEE J Biomed Health Inform*, 2020, 24(6):1686-94.

[6] Wang L, Yang S J, Yang S, et al. Automatic thyroid nodule recognition and diagnosis in ultrasound imaging with the YOLOv2 neural network [J]. *World J Surg Oncol*, 2019, 17(1):12.

[7] Peng S, Liu Y, Lv W, et al. Deep learning-based artificial intelligence model to assist thyroid nodule diagnosis and management: a multicentre diagnostic study [J]. *Lancet Digit Health*, 2021, 3(4):e250-9.

[8] Zhou H, Jin Y H, Dai L, et al. Differential diagnosis of benign and malignant thyroid nodules using deep learning radiomics of thyroid ultrasound images [J]. *Eur J Radiol*, 2020, 127:108992.

[9] 张静漪, 罗燕, 刘加林, 等. 卷积神经网络识别正常及异常甲状腺超声图像的价值 [J]. *四川医学*, 2021, 42(3):305-9.

[10] 乔婷婷, 崔志军, 蔡海东, 等. 基于深度学习构建的甲状腺平面显像智能识别甲状腺功能状态诊断模型 [J]. *中华核医学与分子影像杂志*, 2019, 39(7):403-7.

[11] 李潜, 刘春丽, 郭兰伟, 等. 基于深度学习模型的 S-Detect 技术在甲状腺良恶性结节鉴别诊断中的应用价值 [J]. *郑州大学学报 (医学版)*, 2021, 56(2):285-9.

[12] 武宽, 秦品乐, 柴锐, 等. 基于不同超声成像的甲状腺结节良恶性判别 [J]. *计算机应用*, 2020, 40(1):77-82.

[13] 高源, 邓大同, 刘彧, 等. 2020 版 C-TIRADS 在甲状腺结节诊断和处理中的应用价值 [J]. *安徽医科大学学报*, 2022, 57(3):497-9.

Identification of benign and malignant nodules in thyroid ultrasound images based on deep convolutional neural network

Yao Wenjun^{1,2}, Yin Chaoran³, Zhu Hongqing¹, Jiang Jianmin⁴, Pang Xiaoxi⁵, Sun Yining²

(¹Dept of Radiology, ⁵Dept of Nuclear Medicine, The Second Affiliated Hospital of Anhui Medical University, Hefei 230601; ²Intelligent Mechanization Institute, Hefei Institute of Physical Matter, Chinese Academy of Sciences, Hefei 230031; ³Dept of Ultrasonography, Anhui No. 2 Provincial People's Hospital, Hefei 230011; ⁴School of Electronic and Information Engineering, Anhui University, Hefei 230601)

Abstract Objective To explore the clinical application value of deep convolutional neural network for automatic detection and classification of benign and malignant thyroid nodules ultrasound images. **Methods** A total of 1 012 ultrasound images of thyroid nodules were retrospectively selected and labeled. The YOLOv5 network model was constructed to accurately locate the location of thyroid nodules and automatically trim the area of the nodules. At the same time, a GoogLeNet network model was constructed to classify benign and malignant nodules after reduction. **Results** In the collected data set, the mean accuracy of the target detection network for thyroid nodule location detection was 96.2%. Meanwhile, the sensitivity, specificity, accuracy and AUC of the classification network for benign and malignant nodules were 0.885, 0.822, 0.866 and 0.92 respectively, which were significantly higher than those of the AlexNet model (0.81), VGG model (0.86) and MobileNet model (0.76). **Conclusion** The deep convolutional neural network model has high localization and recognition ability for benign and malignant thyroid nodules in ultrasound images, which is helpful to improve the accuracy of automatic image diagnosis.

Key words thyroid nodule; ultrasound image; deep convolutional neural networks; YOLOv5 network