

基于深度学习的甲状腺超声图像修复算法的初步研究

张敏¹ 倪驰明² 温佳恒² 邓子烨² 徐海珊¹ 楼海亚¹ 潘美¹ 李强¹ 周凌¹

张传菊¹ 凌玉¹ 王娇妮¹ 陈娟萍¹ 王高昂² 李世岩¹

¹浙江大学医学院附属邵逸夫医院超声医学科, 杭州 310016; ²浙江大学伊利诺伊大学厄巴纳香槟校区联合学院, 海宁 314400

张敏和倪驰明对本文有同等贡献

通信作者: 李世岩, Email: lishiyang@zju.edu.cn

【摘要】 目的 探讨基于深度学习对被遮挡的甲状腺超声图像进行修复的可行性。**方法** 回顾性收集自 2020 年 1 月至 2021 年 10 月于浙江大学医学院附属邵逸夫医院采集的甲状腺结节图像共 358 张, 对图片进行随机遮挡后, 使用 DeepFillv2 方法对上述图像被遮挡部分进行修复, 比较修复前后图像的灰度值差异。邀请 6 位不同年资医师 (主任医师、主治医师、住院医师各 2 位) 比较修复前后图片是否存在形态差异, 比较不同医师组判断的正确率及对图像差异的检出率。根据 2020 甲状腺结节超声恶性危险分层中国指南 (C-TIRADS) 对甲状腺结节图像进行超声特征提取 (垂直位、实性、极低回声、可疑微钙化、边缘模糊及彗星尾伪像), 比较修复前后图像中甲状腺结节超声特征的一致性。**结果** 图片修复前后的灰度值均方误差范围为 0.274~0.522, 各组医师的正确率及检出率之间差异具有统计学意义 (均 $P < 0.001$), 总体正确率为 51.95%, 总体检出率为 1.79%, 其中主任医师与住院医师组内的正确率及检出率差异同样具有统计学意义 (均 $P < 0.001$)。图像修复前后图像中各超声特征的一致性均高于 70%, 实性、彗星尾伪像的一致性均高于 90%。**结论** 基于深度学习的甲状腺超声图像修复算法可以有效修复被遮挡的甲状腺超声图像, 同时可以保留甲状腺超声图像特征, 有望扩大深度学习图像数据库规模, 促进深度学习在超声领域的发展。

【关键词】 超声检查; 甲状腺结节; 图像修复; 深度学习

基金项目: 浙江省自然科学基金 (LY20H180005)

DOI: 10.3760/cma.j.cn131148-20221109-00762

Preliminary study on thyroid ultrasound image restoration algorithm based on deep learning

Zhang Min¹, Ni Chiming², Wen Jiaheng², Deng Ziyue², Xu Haishan¹, Lou Haiya¹, Pan Mei¹, Li Qiang¹, Zhou Ling¹, Zhang Chuanju¹, Ling Yu¹, Wang Jiaoni¹, Chen Juanping¹, Wang Gaoang², Li Shiyan¹

¹Department of Ultrasound in Medicine, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou 310016, China; ²Zhejiang University-University of Illinois at Urbana-Champaign Institute, Haining 314400, China

Zhang Min and Ni Chiming contributed equally to this article

Corresponding author: Li Shiyan, Email: lishiyang@zju.edu.cn

【Abstract】 Objective To explore the feasibility of deep learning-based restoration of obscured thyroid ultrasound images. **Methods** A total of 358 images of thyroid nodules were retrospectively collected from January 2020 to October 2021 at Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, and the images were randomly masked and restored using DeepFillv2. The difference in grey values between the images before and after restoration was compared, and 6 sonographers (2 chief physicians, 2 attending physicians, 2 residents) were invited to compare the rate of correctness of judgement and detection of image discrepancies. The ultrasound features of thyroid nodules (solid composition, microcalcifications, markedly hypoechoic, ill-defined or irregular margins, or extrathyroidal extensions, vertical orientation and comet-tail artifact) were extracted according to the Chinese Thyroid Imaging Reporting and Data System (C-TIRADS). The consistency of ultrasound features of thyroid nodules before and after restoration were

compared. **Results** The mean squared error of the images before and after restoration ranged from 0.274 to 0.522, and there were significant differences in the rate of correctness of judgement and detection of image discrepancies between physicians of different groups (all $P < 0.001$). The overall accuracy rate was 51.95%, the overall detection rate was 1.79%, there were significant differences also within the chief physicians and resident groups (all $P < 0.001$). The agreement rate of all ultrasound features of the nodules before and after image restoration was higher than 70%, over 90% agreement rate for features such as solid composition and comet-tail artifact. **Conclusions** The algorithm can effectively repair obscured thyroid ultrasound images while preserving image features, which is expected to expand the deep learning image database, and promote the development of deep learning in the field of ultrasound images.

【Key words】 Ultrasonography; Thyroid nodules; Image restoration; Deep learning

Fund program: Natural Science Foundation of Zhejiang Province (LY20H180005)

DOI:10.3760/cma.j.cn131148-20221109-00762

深度学习在超声影像领域具有巨大的潜力^[1], 目前已经开发出多款人工智能 (artificial intelligence, AI) 辅助诊断系统, 如 AmCAD-UT、S-Detect 等^[2-3], 可以对甲状腺及乳腺结节进行定位及良恶性判断等^[4-6], 但仍无法满足判断疾病预后或肿瘤分期等临床需求。为了构建新的 AI 模型以完成更高阶的临床任务, 需要更多高质量的图像样本。然而, 目前缺乏大规模、高质量、正确标记的超声图像数据集^[7-8]。如果仅依靠个别超声医师采集符合深度学习模型训练要求的图片, 则存在工作量大、时间周期长等阻碍, 无法在短时间内获得足够样本量, 严重影响研究效率。虽然在日常工作中可以采集到大量的超声图像, 但大部分图片均无法满足 AI 模型训练的要求, 其中, 图片中包含的测量及标注痕迹所带来的遮挡问题最为突出, 测量及标注痕迹不仅会遮挡图像中代表病理信息的超声特征, 而且会导致在模型训练过程中目标病变识别和代表性图像的选择发生变化, 引起模型过拟合, 导致模型对病变区域的检出及诊断能力退化^[9]。大量存在标记的图像不仅不能用于开发新的深度学习模型, 同样也无法用于现有模型的自动检测及诊断, 因此如何去除图像中的测量标记, 是需要解决的第一个关键问题。已有研究证明深度学习对于内镜、冠状动脉增强 CT 以及磁共振等图像及视频修复的有效性^[10-12]。本研究拟对甲状腺超声图像进行随机遮挡, 模拟超声图像中的手动标记的遮挡效果, 通过图像修复和补全技术^[13], 建立用于修复被遮挡的甲状腺超声图像的深度学习算法模型, 旨在扩大深度学习训练集的图片数量, 促进深度学习在超声图像相关领域发展。

资料与方法

一、研究对象

回顾性选取 2020 年 1 月至 2021 年 10 月在浙江大学医学院附属邵逸夫医院超声医学科连续采集的甲

状腺结节患者的超声图像共 358 张。纳入标准: 图片能够完整展示至少一个甲状腺结节, 且无任何测量或标注痕迹, 超声图像以 DICOM 或 JPG 格式导出。对图像进行预处理, 包括: 归一化处理, 将图像的分辨率统一到相近大小; 标准化裁剪, 使得图像大小一致。随后仿照人工标记的大小对图像进行遮挡, 并对遮挡大小进行扰动, 模拟超声图像采集过程中各种类型人为标记对甲状腺超声图像的遮挡效果, 为构建图像修复模型提供训练图片集。本研究经医院伦理委员会审批通过 (伦理号: 20190520-70)。

二、仪器与方法

应用 Philips Epiq 5、Philips iU22、GE Logiq E9、Supersonic Aixplorer 超声诊断仪, 分别采用 L12-5、ML6-15 或 SL15-4 线阵探头 (Philips Epiq5 及 Philips iU 22 均使用 L12-5 线阵探头), 频率 4~15 MHz。由已经经过正规培训且正常上岗的超声医师进行甲状腺超声检查并采集结节图像, 流程如下: 患者取去枕仰卧位, 充分暴露颈部, 对甲状腺及双侧颈部进行常规扫查, 识别到甲状腺结节后, 对结节进行横切、纵切扫查并留存甲状腺结节的灰阶超声图像。

三、AI 算法构建

图像修复模型^[13]包含了三个网络架构, 分别为基于编码-解码器的粗粒度修复网络、精粒度修复网络以及基于谱归一化的马尔可夫判别器。基本原理是通过基于深度学习模型的自编码器, 输入损坏或人为标记的图片, 输出修复的估计图片; 修复是通过训练图片样本, 让算法模型学习超声图像的纹理特征, 并尽可能地与周围区域的图像纹理相融合; 基于统计的方式, 估计出缺损位置像素值的大小。此外, 还应用了对抗生成训练方式, 通过一个后置判别器, 使修复的图像尽可能像真实的超声图像。

首先, 将带有掩膜的待重构超声图像输入到粗粒度修复网络中, 网络则生成出模糊的粗粒度重构图像。在该修复网络中, 采用了门控卷积, 代替了以往

的常规卷积操作。门控卷积是一个带有软阈值的筛选操作,对输入有选择机制。该软性选择,可视为按照重要性乘以 0~1 之间的一个权重,作为重构时的参考。基于软性选择的门控卷积可以自动根据梯度更新规则从数据里面学习到软性掩膜的参数。其次,将重构出的模糊图像输入到精粒度修复网络中,该精粒度修复网络与 DeepFillv1^[14]类似,其编码器分为两个平行分支:上下文注意力分支和常规的编码分支。通过两路分支,使模型可借助掩膜周围信息以重构出更精细的图像内容。最后,利用 SN-PatchGAN^[15]作为判别器,使生成的图像更加真实。其中, PatchGAN 输出的是一幅尺寸小于原图的 $N \times N$ 的矩阵,使其每个元素对应原图的一片感受野,可对每一局部特征进行约束。而且 PatchGAN 是全卷积网络,不受输入分辨率的限制。另一方面,利用谱归一化保证神经网络满足 1-lipschitz 连续性,使得参数变化也会更稳定,不容易出现梯度爆炸,保证网络训练更加稳定。

在训练时,结合 L1 范数的重建损失和基于 SN-PatchGAN 的铰链对抗损失进行模型优化。在训练阶段,重构修复网络与判别器交替进行更新。在测试阶段,只用重构修复网络生成最终结果。整个框架结构见图 1。

四、修复效果评价方法

1. 计算图像的灰度均方误差 (mean square error, MSE)

对修复前后图像进行客观评估,将修复前后图像的灰度值范围设置为 0~255,通过对修复前后图像同一位置上的灰度值进行比较判断图像的相似性,即计算修复后图像与原始图像对应位置灰度值之差平方和的均值,即 MSE。MSE 的数值越小,表示修复后的图像越接近原始图像。MSE 的计算公式如下:

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

2. 超声医师读图进行主观评估

A 部分:将入选的 358 张的原始图片及其对应的

修复后图像随机分为三部分,即第一部分将 1 张原始图片复制至 3 张为一组 (共 102 组),第二部分将 1 张修复后图像复制至 3 张为一组 (共 88 组),第三部分将 1 张原始图片复制至 2 张 + 其对应的修复后图像 (共 3 张图片) 为一组 (共 168 组),见图 2。上述各组图片组内及组间均随机乱序排列。分别选择主任医师、主治医师、住院医师各 2 位进行读图,所有医师仅被告知每组的 3 张图片中至少有 2 张是相同的,且每位医师对其他人的读图结果均不知情。需要每位医生回答两个问题:①3 张图片是否完全相同;②若有不同,请选出不同的 1 张。如果未能正确找出第三部分中的修复后图片,或误将 3 张完全相同图片判定为存在不同即为判断错误,其他情况为判断正确,计算随机作答的最高正确率、最高检出率、不同年资医师个体正确率、总体正确率、个体检出率及总体检出率。

随机最高正确率 = (第一部分 102 组 + 第二部分 88 组) / 358 组 $\times 100\%$ = 53.07%;

随机最高检出率 = 33.3% (3 张图片随机选择 1 张,检出的正确率为 33.3%);

个体正确率 = 医师个体判断正确的组数 / 358 组 $\times 100\%$;

总体正确率 = 6 位医师均判断正确的组数 / 358 组 $\times 100\%$;

个体检出率 = 医师个体正确检出差异图像的组数 / 168 组 $\times 100\%$;

总体检出率 = 6 位医师均正确检出差异图像的组数 / 168 组 $\times 100\%$ 。

B 部分:将原始图片随机等分为两部分,使用第一部分的原始图片以及第二部分原始图像对应的修复后图像组成图片库 1,再由第二部分的原始图像以及第一部分原始图片对应的修复后图像组成图片库 2,见图 3,两个图片库中图片乱序排列。由另外 3 位不同年资超声医师 (主任医师、主治医师、住院医师) 分别根据 C-TIRADS^[16]对图像中甲状腺结节的超声

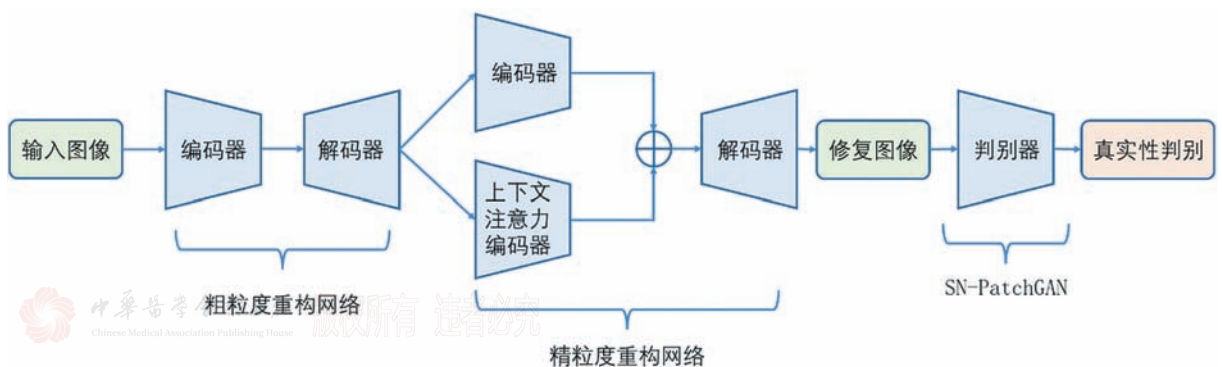


图 1 图像修复算法构建的框架结构示意图

特征进行分析，部分结节图像中未包含颈前肌群，此类结节回声特征判定为待定，每位医师对其他人的读图结果均不知情。先阅读图片库 1，1 个月后再阅读图片库 2，最后计算 3 位医师对修复前后图像超声特征判读的一致性。

五、统计学方法

采用 SPSS 25.0 统计软件分析数据。计数资料以率表示，应用卡方检验比较不同年资医师的个体正确率、个体检出率及其与随机最高正确率和随机最高检出率之间的差异。以双侧 $P < 0.05$ 为差异有统计学意义。

结 果

一、图片样本情况

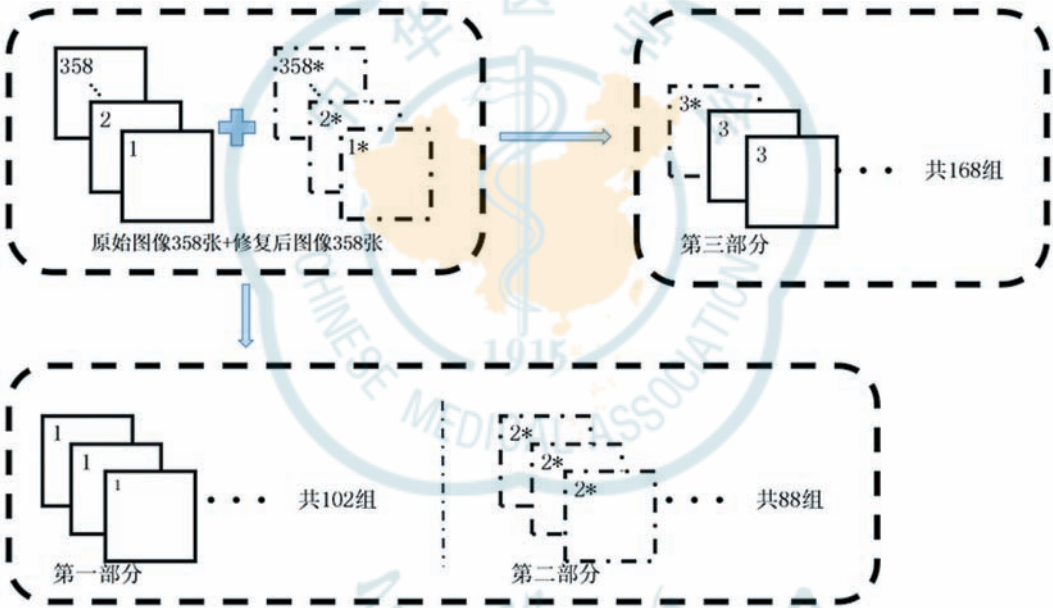


图 2 A 部分图库组成示意图

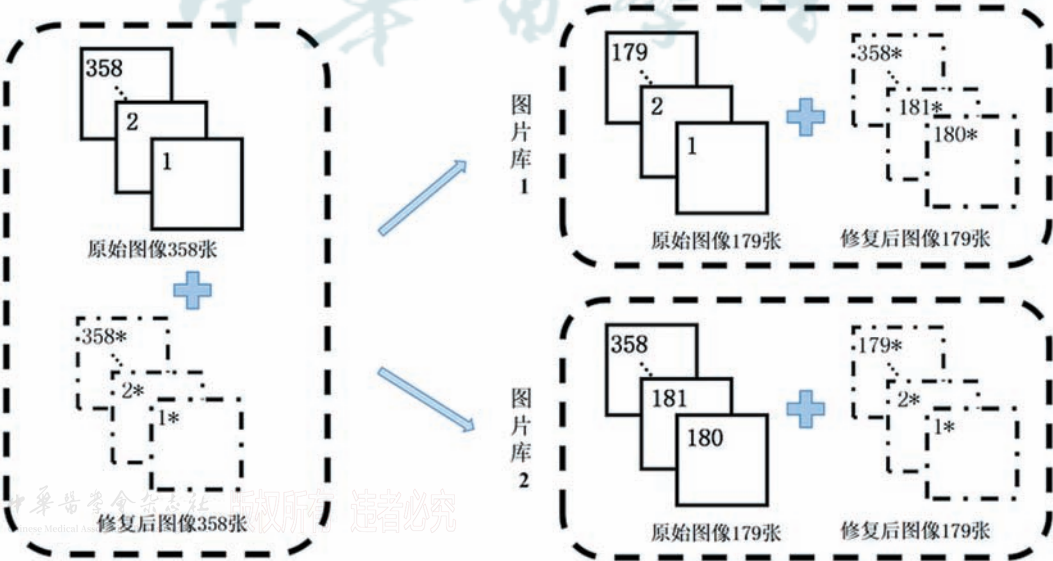


图 3 B 部分图库组成示意图

对 358 张图像进行人工遮挡处理，得到待修复图像，随后使用 DeepFillv2 方法对图像进行修复，见图 4。

二、图像的灰度值 MSE

对修复前后图像灰度值差异进行比较，计算所得的 MSE 为 0.274~0.522，人眼在不同灰度值背景下可分辨的灰度值差异一般在 10 左右^[17]，该修复结果远小于人眼可感知的灰度值差异。

三、超声医师读图结果

A 部分：

1. 正确率分析

6 位医师判断正确的图片组合分别为 238 组、310 组、204 组、205 组、323 组、221 组，个体正确率分别为 66.48%、86.59%、56.98%、57.26%、

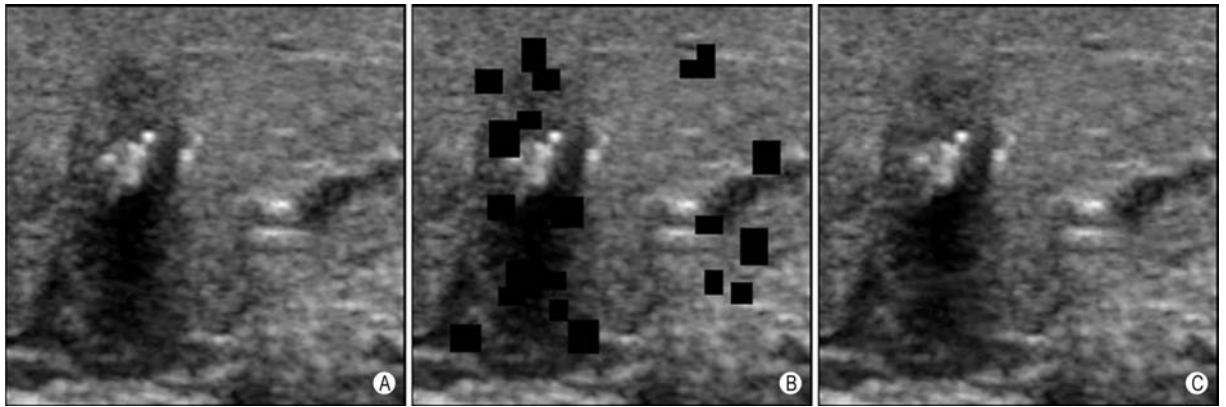


图 4 甲状腺结节图片遮挡及修复 A: 原始超声图像; B: 进行随机遮挡(黑色方框)后的图像; C: 经图像修复算法修复后的图像

90.22%、61.73%，见图 5。不同医师之间正确率差异有统计学意义 ($P < 0.001$)，且正确率并不随着医师年资增长而升高。

主任医师组平均正确率为 75.98%，主治医师组平均正确率为 57.12%，住院医师组平均正确率为 76.54%，对各组医师的读图正确率进行组内比较，主任医师组内及住院医师组内差异有统计学意义 (均 $P < 0.001$)，而主治医师组内差异无统计学意义 ($P = 0.940$)。

6 位医师均判断正确的组数为 186 组，总体正确率为 51.95%，低于随机最高正确率，但与随机最高正确率之间差异无统计学意义 ($P = 0.765$)。

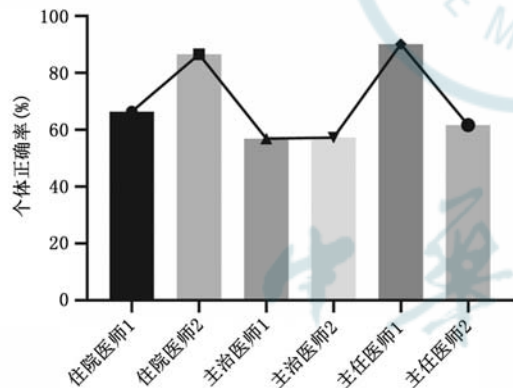


图 5 不同医师的个体正确率

2. 检出率分析

不同医师对第三部分图像的检出率分别为 28.57%、71.43%、8.33%、8.93%、79.17%、18.45%，见图 6。不同医师之间检出率差异有统计学意义 ($P < 0.001$)，且检出率并未随着医师年资增长而升高。6 位医师均正确检出的组数仅为 3 组，总体检出率为 1.79%，显著低于随机最高检出率 (33.3%) ($P < 0.001$)。

B 部分：3 位医师对原始图像及修复后图像进行读片，判断图像中结节的超声特征，两组图片修复前后的超声特征判断一致性见表 1，所有超声特征的一

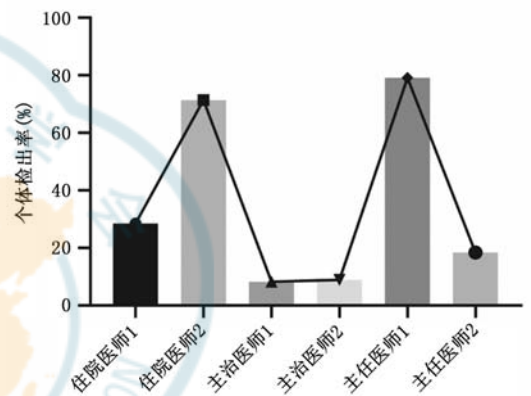


图 6 不同医师对第三部分图像的个体检出率

致性均高于 70%，其中对于实性及彗星尾伪像的判断一致性均高于 90%。

表 1 各年资医师对图像修复前后的甲状腺结节超声特征判断一致性 (%)

不同医师	垂直位	实性	极低回声	可疑微钙化	边缘模糊	彗星尾伪像
主任医师	84.08	92.74	77.09	75.14	71.79	94.69
主治医师	86.59	94.41	76.26	84.91	80.72	99.16
住院医师	84.35	98.89	85.19	77.10	79.33	98.88

讨 论

超声诊断水平高度依赖于操作者^[18]，由于不同医师的经验和诊疗水平不同，对病变的检出率及诊断准确性存在差异。而经过训练后的 AI 模型可以对超声医师肉眼不易感知的像素级差异进行分析，对病变进行定位及良恶性判断^[19]，且具有较好的可重复性^[20]，能够辅助超声医师进行诊断。稳健的模型通常需要大量样本进行训练，由于目前患者的隐私和医疗机构之间存在的信息流通限制，导致创建多中心大样本的大型数据集较为困难^[8]。有很多方法可以解决模型训练过程中样本量不足的问题：迁移学习通过将权重信息从源域转移到目标域以获得新的预测模型，但是所取得的模型效果不及基于大型图片库所取得的

模型^[21], 会影响算法的泛化能力^[22]; 数据增强技术通过使用图像旋转、图像翻转、图像裁剪等多种组合以扩大图片集的规模^[23-24], 然而该技术无法将新的特征引入模型, 只能针对原有特征进行强化训练, 影响模型对图像信息的进一步深入挖掘^[25]。为了能够训练出既稳定又可以实际应用的 AI 模型, 创建大样本数据集仍然是非常有必要的。

虽然日常工作中可以获得大量超声图像, 但是大部分图像无法直接用于深度学习模型的训练。究其原因, 主要有以下几个方面: ①由于测量和标注需要, 图片中往往存在手动标注的痕迹, 可能扰乱深度学习模型的判断机制^[26]; ②图片增益调节不当, 或者是因为在传输过程中产生偏差, 导致图像过亮或过暗; ③深度及靶目标比例调节不当, 导致感兴趣的区域显示不全或过小, 显示不全将影响对病灶特征的完整提取, 而显示过小会导致病灶局部有效分辨率降低, 特征信息丢失; ④超声医师对于病变性质的判断存在一定的主观性^[27], 诊断错误的超声图像会在训练过程中产生标签噪声, 需要病理结果证实超声医师诊断的正确性, 而日常工作中储存的大部分图片并无对应的病理结果, 导致符合要求的图片数量进一步减少。如何将大量不符合标准的超声图像用于深度学习模型训练, 以提高模型的诊断效能和鲁棒性, 是亟需解决的问题; 其中如何去除图像中的测量标记, 则是需要解决的第一个关键问题。本研究使用随机遮挡模拟图像中标记的遮挡效果, 随后使用 DeepFillv2 方法对甲状腺超声图像进行修复, 建立了一种稳定的超声图像修复模型, 用于去除超声图像中的标记。

目前存在多种图像标记检出及去除算法。Li 等^[28]使用瑞利最大似然切换双边滤波器的方法检测图像中的噪声及斑点。Sun 等^[29]使用基于边缘连接和 Criminisi 图像恢复算法进行超声图像中的标记去除。此外, 还有结合多块卷积变分信息瓶颈模型的图像修复模型、甲状腺超声图像级联标记物去除算法模型, 可以对图像中的标记进行有效的识别及去除^[9,26]。虽然目前已经存在成熟的图像修复算法模型, 但对于医学超声图像修复的研究仍较少, 且大部分关于超声影像的研究仅论证了算法去除标记的可行性, 并未论证算法修复过程是否会对超声图像所包含的超声特征造成影响。本算法通过对抗生成的策略, 使修复的图像比其他利用人为规则方式生成的图像更加真实; 整个深度网络可端到端地学习, 避免了传统方法中经验性的人为干预; 在验证算法去除标记的有效性之外, 对超声图像中代表病理信息的超声特征进行评估, 使得算法更贴近临床需求。

在本研究中, 首先对修复前后的图像进行客观评

价, 分析修复前后图片对应位置的灰度值差异, 其灰度值 MSE 范围为 0.274~0.522, 该差异远小于人眼可感知的最小灰度差异^[17], 即修复前后图像同一位置上的灰度值差异低于人眼的识别范围。在客观评价方面, 本算法对图像具有较好的修复能力。

其次, 通过医师对图像差异的识别进行主观评价。在 A 部分图库的生成过程中半数左右的相同图片组与存在差异的图片组为混杂排布, 减少医师在读图过程中随机作答对结果的影响。结果表明医师的正确率及检出率并不随着医师的年资增长而升高, 显示医师对于图片差异的检出能力可能与经验无关; 排除因医师随机作答、医师经验差异等可能影响正确率的因素, 进一步比较医师组内的正确率, 可以发现主治医师组组内正确率差异无统计学意义, 其正确率接近随机最高值, 主任医师组及住院医师组正确率高于随机最高值, 但是这两组组内的正确率差异有统计学意义, 该差异可能是由于参与读图医师人数较少而导致的抽样误差。为减少抽样误差对结果的影响, 本研究计算了总体正确率, 总体正确率 (51.95%) 与随机最高正确率 (53.07%) 的差异并无统计学意义。检出率在主任医师组及住院医师组差异同样有统计学意义, 且总体检出率 (1.79%) 显著低于随机最高检出率 (33.3%)。对医师的正确率及检出率分析发现, 医师对于图像差异的判断能力与随机作答的结果相当甚至不及, 表明修复前后的图像在医师的主观判断中差异较小, 修复后的图像在一定程度上可以起到以假乱真的效果。在客观评价及医师主观评价中, 本算法对于超声图像均具有较好的修复效果。

在临床诊断过程中, 甲状腺结节超声特征直接影响 C-TIRADS 分级及患者后续治疗方案的制定。为了进一步探讨修复过程是否会影响图像中甲状腺结节的超声病理特征及诊断结果, 本研究根据 C-TIRADS^[16]对修复前后图像的超声特征进行评估。de Vet 等^[30]在对 Cohen's Kappa 一致性检验的研究提出, 随着结局变量中某一变量的占比升高导致样本分布不均衡时, 会出现 Kappa 值低, 而实际一致性较高的情况, 此时使用判断一致的比例进行评价更加可靠。本研究中各超声特征样本分布不均衡, 如结节大部分为实性等, 故本研究使用图像修复前后各超声特征的一致性比例进行一致性评估, 结果显示不同年资医师对于图像修复前后结节垂直位、可疑微钙化、极低回声及边缘模糊的判断中一致性均高于 70%, 其中对于实性以及彗星尾伪像判断的一致性均在 90% 以上。该结果表明本算法修复前后图像的超声特征具有较好的一致性, 证实本研究的算法对图像进行有效修复的同时可以保留图像中代表病理信息的超声

特征,避免因修复过程导致的信息丢失引起误诊漏诊。此外,本研究中纳入了多种品牌的超声仪所采集的超声图像,同时使用随机遮挡模拟多种标记对于图像的遮挡效果,增加了训练集图像的多样性,使本算法的泛化能力进一步提高。未来有望将大量闲置的不同来源的超声图片从非标准采集影像转换为可供深度学习模型训练的理想图像,通过“变废为宝”的转换,可以使样本量上升一到两个数量级,在很大程度上提高各个医疗中心超声影像数据库的 AI 研究能力,从而更好地解决临床难题。

本研究仍存在一定的局限性:①目前仅局限于甲状腺结节超声图像的修复,未来尝试将该修复方法应用于其他器官的超声图像,以拓展本算法的适用范围;②本研究的样本量较少,未来将根据需要对本修复算法进行更大样本量的训练,以提升算法的稳定性和修复能力;③本研究仅通过图像的灰度值 MSE 以及医师肉眼读图的方式验证算法修复的效果,未来仍需通过比较原始图像与修复图像在其他深度学习模型中应用的差异进一步验证该修复模型的效果。

综上所述,本研究提出了一种利用对抗生成策略的图像修复算法,该算法结合粗粒度和精粒度编码-解码器,可渐进式地重构出高清图像,结合判别器对抗学习,使修复图像更加真实。该方法可以有效对被遮挡的图像进行修复,同时保留图像原有的超声特征,未来有望用于非标准采集超声影像的标准化处理,扩大深度学习图像数据库的规模,促进深度学习在超声领域的发展。

利益冲突 所有作者均声明不存在利益冲突

作者贡献声明 张敏、倪驰明:酝酿和设计实验,实施研究,分析/解释数据,起草文章,对文章的知识性内容作批评性审阅;温佳恒、邓子烨、徐海珊、楼海亚、潘美、李强、周凌、张传菊、凌玉、王娇妮、陈娟萍:实施研究,分析/解释数据,对文章的知识性内容作批评性审阅;王高昂、李世岩:酝酿和设计实验,实施研究,分析/解释数据,对文章的知识性内容作批评性审阅

参 考 文 献

- [1] Makimoto H, Hökmann M, Lin T, et al. Performance of a convolutional neural network derived from an ECG database in recognizing myocardial infarction [J]. *Sci Rep*, 2020, 10 (1): 8445. DOI: 10. 1038/s41598-020-65105-x.
- [2] 徐锦媚,刘萌,王鑫毅,等.计算机辅助诊断系统联合弹性成像及高年资超声医师判读鉴别甲状腺结节良恶性的价值[J]. *临床超声医学杂志*, 2022, 24 (1): 26-30. DOI: 10. 16245/j. cnki. issn1008-6978. 2022. 01. 012.
- [3] 王琳萍,张波.计算机辅助诊断系统在甲状腺结节超声诊断中的应用进展[J]. *医学研究杂志*, 2021, 50 (10): 158-161. DOI: 10. 11969/j. issn. 1673-548X. 2021. 10. 035.
- [4] Zhou H, Jin Y, 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. DOI: 10. 1016/j. ejrad. 2020. 108992.
- [5] Zhu J, Zhang S, Yu R, et al. An efficient deep convolutional neural network model for visual localization and automatic diagnosis of thyroid nodules on ultrasound images [J]. *Quant Imaging Med Surg*, 2021, 11 (4): 1368-1380. DOI: 10. 21037/qims-20-538.
- [6] Lingam RK, Qarib MH, Tolley NS. Evaluating thyroid nodules: predicting and selecting malignant nodules for fine-needle aspiration (FNA) cytology [J]. *Insights Imaging*, 2013, 4 (5): 617-624. DOI: 10. 1007/s13244-013-0256-6.
- [7] Ching T, Himmelstein DS, Beaulieu-Jones BK, et al. Opportunities and obstacles for deep learning in biology and medicine [J]. *J R Soc Interface*, 2018, 15 (141): 20170387. DOI: 10. 1098/rsif. 2017. 0387.
- [8] Karimi D, Dou H, Warfield SK, et al. Deep learning with noisy labels: exploring techniques and remedies in medical image analysis [J]. *Med Image Anal*, 2020, 65: 101759. DOI: 10. 1016/j. media. 2020. 101759.
- [9] Yao S, Yan J, Wu M, et al. Texture synthesis based thyroid nodule detection from medical ultrasound images: interpreting and suppressing the adversarial effect of in-place manual annotation [J]. *Front Bioeng Biotechnol*, 2020, 8: 599. DOI: 10. 3389/fbioe. 2020. 00599.
- [10] Ali S, Zhou F, Bailey A, et al. A deep learning framework for quality assessment and restoration in video endoscopy [J]. *Med Image Anal*, 2021, 68: 101900. DOI: 10. 1016/j. media. 2020. 101900.
- [11] Tatsugami F, Higaki T, Nakamura Y, et al. Deep learning-based image restoration algorithm for coronary CT angiography [J]. *Eur Radiol*, 2019, 29 (10): 5322-5329. DOI: 10. 1007/s00330-019-06183-y.
- [12] Eun DI, Jang R, Ha WS, et al. Deep-learning-based image quality enhancement of compressed sensing magnetic resonance imaging of vessel wall: comparison of self-supervised and unsupervised approaches [J]. *Sci Rep*, 2020, 10 (1): 13950. DOI: 10. 1038/s41598-020-69932-w.
- [13] Yu J, Lin Z, Yang J, et al. Free-Form Image Inpainting With Gated Convolution [C]. *Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019. DOI: 10. 48550/arXiv. 1806. 03589.
- [14] Yu J, Lin Z, Yang J, et al. Generative Image Inpainting with Contextual Attention [C]. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018. DOI: 10. 48550/arXiv. 1801. 07892.
- [15] Isola P, Zhu JY, Zhou T, et al. Image-to-Image Translation with Conditional Adversarial Networks [C]. *Proceedings of the IEEE Conference on Computer Vision & Pattern Recognition*, 2016. DOI: 10. 48550/arXiv. 1611. 07004.

- [16] 中华医学会超声医学分会浅表器官和血管学组, 中国甲状腺与乳腺超声人工智能联盟. 2020 甲状腺结节超声恶性危险分层中国指南: C-TIRADS[J]. 中华超声影像学杂志, 2021, 30(3): 185-200. DOI: 10.3760/cma.j.cn131148-20210205-00092. Superficial Organ and Vascular Ultrasound Group, Society of Ultrasound in Medicine, Chinese Medical Association; Chinese Artificial Intelligence Alliance for Thyroid and Breast Ultrasound. 2020 Chinese Guidelines for Ultrasound Malignancy Risk Stratification of Thyroid Nodules: The C-TIRADS[J]. Chin J Ultrasonogr, 2021, 30(3): 185-200. DOI: 10.3760/cma.j.cn131148-20210205-00092.
- [17] 范晓鹏, 蔡铁峰, 朱枫. 实现人眼灰度感知信息最大化的图像增强算法[J]. 计算机辅助设计与图形学学报, 2015, 27(10): 1900-1906. Fan XP, Cai TF, Zhu F. Image enhancement algorithm for achieving maximum of perceived grayscale information of human visual system [J]. Journal of Computer-Aided Design & Computer Graphics, 2015, 27(10): 1900-1906.
- [18] Chambara N, Ying M. The diagnostic efficiency of ultrasound computer-aided diagnosis in differentiating thyroid nodules: a systematic review and narrative synthesis[J]. Cancers (Basel), 2019, 11(11): 1759. DOI: 10.3390/cancers11111759.
- [19] Hosny A, Parmar C, Quackenbush J, et al. Artificial intelligence in radiology[J]. Nat Rev Cancer, 2018, 18(8): 500-510. DOI: 10.1038/s41568-018-0016-5.
- [20] 于腾飞, 何文, 甘从贵, 等. 基于深度学习超声在乳腺肿块四分类中的应用价值[J]. 中华超声影像学杂志, 2020, 29(4): 337-342. DOI: 10.3760/cma.j.cn131148-20190828-00519. Yu TF, He W, Gan CG, et al. Application value of deep learning ultrasound in the four-category classification of breast masses [J]. Chin J Ultrasonogr, 2020, 29(4): 337-342. DOI: 10.3760/cma.j.cn131148-20190828-00519.
- [21] Kermany DS, Goldbaum M, Cai W, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning[J]. Cell, 2018, 172(5): 1122-1131. e9. DOI: 10.1016/j.cell.2018.02.010.
- [22] Bauer Z, Gomez-Donoso F, Cruz E, et al. UASOL, a large-scale high-resolution outdoor stereo dataset[J]. Sci Data, 2019, 6(1): 162. DOI: 10.1038/s41597-019-0168-5.
- [23] Soffer S, Ben-Cohen A, Shimon O, et al. Convolutional neural networks for radiologic images: a radiologist's guide [J]. Radiology, 2019, 290(3): 590-606. DOI: 10.1148/radiol.2018180547.
- [24] Chi J, Walia E, Babyn P, et al. Thyroid nodule classification in ultrasound images by fine-tuning deep convolutional neural network[J]. J Digit Imaging, 2017, 30(4): 477-486. DOI: 10.1007/s10278-017-9997-y.
- [25] Xie X, Niu J, Liu X, et al. A survey on incorporating domain knowledge into deep learning for medical image analysis[J]. Med Image Anal, 2021, 69: 101985. DOI: 10.1016/j.media.2021.101985.
- [26] Ying X, Zhang Y, Yu M, et al. Cascade marker removal algorithm for thyroid ultrasound images [J]. Med Biol Eng Comput, 2020, 58(11): 2641-2656. DOI: 10.1007/s11517-020-02216-7.
- [27] Nir G, Hor S, Karimi D, et al. Automatic grading of prostate cancer in digitized histopathology images: learning from multiple experts[J]. Med Image Anal, 2018, 50: 167-180. DOI: 10.1016/j.media.2018.09.005.
- [28] Li H, Wu J, Miao A, et al. Rayleigh-maximum-likelihood bilateral filter for ultrasound image enhancement [J]. Biomed Eng Online, 2017, 16(1): 46. DOI: 10.1186/s12938-017-0336-9.
- [29] Sun M, Meng Q, Wang T, et al. Removal of manually induced artifacts in ultrasound images of thyroid nodules based on edge-connection and criminisi image restoration algorithm[J]. Comput Methods Programs Biomed, 2021, 200: 105868. DOI: 10.1016/j.cmpb.2020.105868.
- [30] de Vet HC, Mookink LB, Terwee CB, et al. Clinicians are right not to like Cohen's κ [J]. BMJ, 2013, 346: f2125. DOI: 10.1136/bmj.f2125.

(收稿日期: 2022-11-09)

