

基于残差注意力机制的肺结节数据增强方法



李 阳^{1,2}, 李春璇^{2*}, 徐灿飞³, 方立梅⁴

(1. 东北师范大学前沿交叉研究院 长春 130000; 2. 长春工业大学计算机科学与工程学院 长春 130012;
3. 浙江大学医学院附属邵逸夫医院 杭州 310000; 4. 长春中医药大学附属第三附属医院 长春 130117)

【摘要】针对带标注的肺 CT 图像数据匮乏而导致的深度学习模型训练困难,以及现有生成算法生成肺结节不同特征模糊、细节丢失的问题,提出了肺结节图像的数据增强 RAU-GAN 算法。首先,在生成器网络中嵌入残差注意力模块,该模块可以聚焦于局部不同的感兴趣区域,以实现肺结节与背景信息的独立生成,并且重新设计了注意力模块中的残差块来减少网络的深度和训练的复杂度。其次,将判别器设计为 U-Net 架构,可以给更新后的生成器反馈更多信息,以提高判别性能。最后,在数据集 LUNA16 和 Deep Lesion 上进行实验,结果与现有方法相比,在视觉效果和不同评价指标上均有提升,验证了生成图像包含了更丰富的细节信息。

关键词 数据增强; Pix2Pix; RAU-GAN; 残差注意力机制; U-Net 判别器
中图分类号 TP391.4 **文献标志码** A **doi**:10.12178/1001-0548.2022363

Data Augmentation of Lung Nodule Based on Residual Attention Mechanism

LI Yang^{1,2}, LI Chunxuan^{2*}, XU Canfei³, and FANG Limei⁴

(1. Academy for Advanced Interdisciplinary Studies, Northeast Normal University Changchun 130000;
2. School of Computer Science and Engineering, Changchun University of Technology Changchun 130012;
3. Run Run Shaw Hospital, Zhejiang University School of Medicine Hangzhou 310000;
4. The Third Affiliated Clinical Hospital of Changchun, University of Traditional Chinese Medicine Changchun 130117)

Abstract Aiming at the difficulty of deep learning model training caused by the lack of labeled lung Computed Tomography (CT) image data and the lung nodule feature model generated by existing generation algorithms. To solve the problem of blur and detail loss, a data-enhanced RAU-GAN algorithm for pulmonary nodule images is proposed. Firstly, a residual attention module is embedded in the generator network, which can focus on different local regions of interest to achieve the independent generation of lung nodules and background information. Moreover, the residual block structure in the attention module is redesigned to reduce the depth of the network and training complexity. Second, the discriminator is designed as U-Net architecture, which can feed back more information to the updated generator to improve the discrimination performance. Finally, experiments were conducted on data set LUNA16 and deep lesion. The results show that the visual and different evaluation indexes have improved in comparison with existing methods, which verifies that the generated images can contain richer details.

Key words data augmentation; Pix2Pix; RAU-GAN; residual Attention mechanism; U-Net discrimination

癌症已成为人口死亡的主要原因之一,而肺癌的发病率和致死率均较高^[1]。早期肺癌在胸部计算机断层扫描 (computed tomography, CT) 上的表现为

肺结节,对肺结节的早期检测可以提高肺癌患者的存活率。计算机辅助检测 (computer aided detection, CAD) 技术能够辅助医生进行医学检测,减少了医

收稿日期: 2022-10-23; 修回日期: 2023-10-24

基金项目: 国家自然科学基金 (NSFC 61806024, NSFC12226003, NSFC12171076); 吉林省科技厅重点研发计划 (20210201081GX, 20200401103GX); 吉林省教育厅重点科研项目 (JKH20220685KJ, JKH20220692KJ); 吉林省卫生厅卫生科研人才专项 (2020SCZ25)

作者简介: 李阳 (1979-), 女, 博士, 教授, 主要从事机器学习、图像处理方面的研究。

*通信作者: 李春璇, E-mail: 2202003054@stu.ccut.edu.cn

生工作量的同时也提高了诊断的准确性。近年来深度学习在肺 CAD 诊断方面得到广泛应用。然而深度学习模型的训练必须依赖大量数据, 实际中由于人工标注成本昂贵导致获取大量带标注的肺 CT 影像非常困难, 数据增强技术为数据的扩充提供了可能。

数据增强是一种数据扩充技术, 传统数据增强方法有旋转、平移、裁剪、缩放、噪声扰动等。这些方法被广泛应用于扩充训练集^[2-3], 然而传统数据增强方法的输出过度依赖于原始数据, 深度模型利用扩充数据来训练模型容易出现过拟合问题。近年来, 许多研究使用生成式对抗网络 (generative adversarial networks, GAN) 进行数据扩充。GAN 通过两个网络的对抗博弈过程来学习真实样本的数据分布^[4], 与传统数据增强方法对比, 本质上提高了样本特征的多样性。

目前 GAN 衍生出一系列的变体模型, 改进的变体模型逐步从自然场景的应用跨越到医学图像的应用^[5-6]。在肺结节的生成任务中, 文献 [7] 为了提高分割网络训练模型性能、增加训练数据, 提出基于 Style-GAN 的新型数据增强方法合成肺结节, 先从整体中提取样式和语义标签, 然后用随机选择的样式为每个语义标签合成增强的 CT 图像。文献 [8] 为提高检测网络的性能, 对肺结节数据集进行增强, 提出一种基于计算机断层扫描的生成式对抗网络的数据增强方法, 可以在指定位置添加肺结节, 并引入 DropBlock 来解决过拟合的问题。

文献 [9] 提出一种基于 Conditional GAN (CGAN) 的肺结节图像生成算法 CT-GAN, 该网络能学习图像到图像的映射关系, 通过篡改原始的 CT 图像数据, 得到近似真实的医学图像数据, 从而扩充正样本数据。

在 GAN 的增强任务中, 引入注意力模块可以达到更优的生成效果。在生成式对抗网络的生成任务中最常用的注意力模块是自注意力机制和通道注意力机制。SAGAN^[10] 在生成器和判别器中都添加了自注意力模块, 在每一层都考虑了全局信息, 在提高感受野和减小参数量之间找到了一个很好的平衡, 生成与全局相关性比较高的图片。文献 [11] 在 CycleGAN 的判别器中添加了空间注意力, 将注意力图反馈到生成器, 来协助生成器可以关注到图像中有区分度的区域, 由此带来了模型性能的持续提升。

U-Net^[12] 被应用在医学图像的初衷是为了解决医学图像分割的问题, 其 U 形结构启发了生成算

法。文献 [13] 提出了 U-Net 的判别器架构, 鼓励 U-Net 判别器更多关注真实图像与伪图像之间的语义和结构变化, 使生成器能生成图片保持全局和局部真实感。文献 [14] 提出一种端到端注意力增强的串行 U-Net++ 网络, 串行 U-Net++ 模块提取不同分辨率的特征并在不同的尺度上重建图像。该模块直接将浅层的原始信息传递给更深层次, 使更深层次专注于残差学习, 而重用浅层上下文信息。

传统的 GAN 网络生成图像效果有限, 训练不稳定且训练过程容易模式崩溃^[15]。直接用 GAN 生成肺结节, 容易存在病灶模糊和背景噪点多的问题, 为解决以上问题, 本文提出利用改进的 Pix2Pix^[16] 模块生成肺结节图像, 主要贡献如下。

1) 生成器中添加设计后的残差注意力模块^[17]。目的是在图像的生成中不但关注到肺结节的生成, 同时也关注到复杂的背景信息。对不同的信息进行特征筛选, 自适应地学习肺结节图像保证图像不同特征细节的生成。

2) 设计残差块结构。残差注意力模块的添加使整个生成网络层数扩展到很深, 重新设计后的残差块减少了网络深度和训练的复杂度, 同时更好地适应生成网络。

3) 设计 U-Net 判别器代替 Pix2Pix 中的马尔可夫判别器。由于 U-Net 的编码器和解码器对应不同模块之间的跳跃链接, 输出层的通道就包含了不同级别的信息, 可以反馈给生成器更详细的信息。

1 算法结构设计

1.1 Pix2Pix 算法原理

GAN 相关的衍生模型越来越多, 在 GAN 中加入各种的条件信息可以更好地控制生成图像的内容和效果, Pix2Pix 是衍生的条件模型之一, 可以达到图像风格转换的目的。Pix2Pix 是基于 CGAN 实现两个域之间的转换模型, 将输入图片作为条件信息, 学习从源域图像到目标域图像之间的映射来生成指定的输出图像。生成器采用的 U-Net 网络能够充分融合特征, 判别器采用的 PatchGAN 分类器能够只在图像块的尺度上进行惩罚, 对输入图像的每一个区域都输出一个预测概率值。其中 Pix2Pix 的损失函数包含了 CGAN 的损失函数和 L1 正则项两部分。

CGAN 的生成器要根据约束条去生成图片, 判别器除了判别生成图像是否为真, 还要判别生成图像与真实图片是否匹配, 优化目标函数可定义为:

$$\mathcal{L}_{cGAN}(G,D) = E_{x,y}[\log_2 D(x,y)] + E_{x,z}[\log_2(1 - D(x,G(x,z)))] \quad (1)$$

式中, G 为生成器; D 为判别器; x 为真实图像; y 为经过预训练后的约束条件; z 为随机噪声。对于 D 而言, 需要最大化目标函数, D 给真实图像高分, 让 $D(x,y)$ 越大越好, 同时给约束条件下生成的图像低分, $D(x,G(x,z))$ 越小越好。对于 G 而言, 需要最小化目标函数, 让 D 给在约束条件下生成的图片高分, $D(x,G(x,z))$ 越大越好。

L1 损失用来约束生成图像与真实图像之间的距离, 可定义为:

$$\mathcal{L}_{L1}(G) = E_{x,y,z}[\|y - G(x,z)\|_1] \quad (2)$$

式中, 矩阵 1-范数表示向量中元素绝对值之和, 即约束条件 y 与生成的图像两者绝对值之和。

最终 Pix2Pix 的目标函数为:

$$G^* = \arg \min_G \max_D \mathcal{L}_{cGAN}(G,D) + \lambda \mathcal{L}_{L1}(G) \quad (3)$$

式中, λ 为超参数。

1.2 RAU-GAN 网络设计

1.2.1 生成器网络

RAU-GAN 将改进后的残差注意力机制与

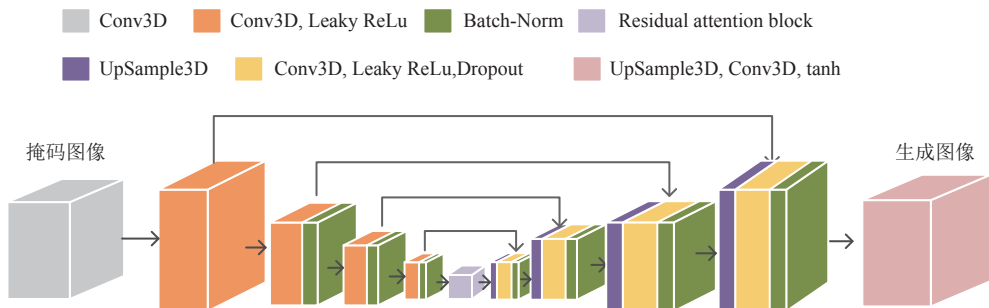


图 1 生成器网络模型

1.2.2 残差块注意力机制

本文采用了残差注意力模块, 结合了通道和空间注意力机制, 可以选择重要的对象和区域。通过残差注意力堆叠式的网络结构来同时关注多个不同的感兴趣区, 堆叠的方式减少了模型的复杂度。通过注意力残差学习来优化残差注意力网络, 避免了网络层数太深引起的梯度消失问题。本文残差注意力由 3 个注意力模块堆叠而成, 每个注意力模块由主干分支和掩膜分支两个部分组成, 如图 2 所示。主干分支进行特征处理, 如结节形状是否存在分叶和毛刺等细节信息。掩膜分支是下采样和上采样的过程, 用来获取特征图的全局信息, 如结节信息。掩膜分支学习得到与主干的输出大小一致的掩膜,

Pix2Pix 生成器相结合。生成器中图像到图像的转换是高分辨率的输入映射到另一个高分辨率的输出, 输入和输出需要粗略对齐, 同时需要共享高层语义信息和底层的语义信息, 故采用 U-Net 全卷积网络。生成器通过左侧不断下采样到达中间的隐含编码层, 再通过右侧上采样来还原图像, 左侧与右侧中间添加跳跃链接将部分有用的重复信息直接共享到生成器中。同时, 添加 L1 损失函数, 来约束生成图像与原始图像之间的特征, 对网络进行优化。

RAU-GAN 的生成器网络结构如图 1 所示, 采用改进后的残差注意力模块嵌入到 U-Net 模块中, 输入掩膜处理后的大小为 $32 \times 32 \times 32$ 的图片, 首先使用大小为 $4 \times 4 \times 4$ 的卷积核, 步幅为 2, 对输入的图像进行处理。再依次通过 4 个卷积层进行特征提取, 之后连接残差注意力模块去更多地关注感兴趣区。同时采用跳跃链接层连接结构, 在对称结构中加入 Dropout 和 Batch-Norm, 可以保留更多的图像细节, 协助反卷积层完成图像的恢复工作, 并且减少梯度消失, 加快模型训练。

通过对特征图的处理输出维度一致的注意力特征图。最后用点乘操作将两个分支的特征图结合得到最终的输出特征图。

此外, 由于残差注意力机制堆叠了多个注意力模块, 每个注意力模块又包含大量的残差块, 因此网络很容易扩展成深层网络。复杂的层次结构添加到生成器, 容易引起生成网络的模式崩溃问题, 为优化网络结构和训练时间, 重新设计了残差块, 其结构如图 3 所示, 新设计的残差块仅由一组批处理归一化 (BN) 层、激活 (ReLU) 层和卷积层 (Conv) 组合而成, 改进后的残差注意力模型网络层减轻的同时也达到了生成注意力感知特征的目的。

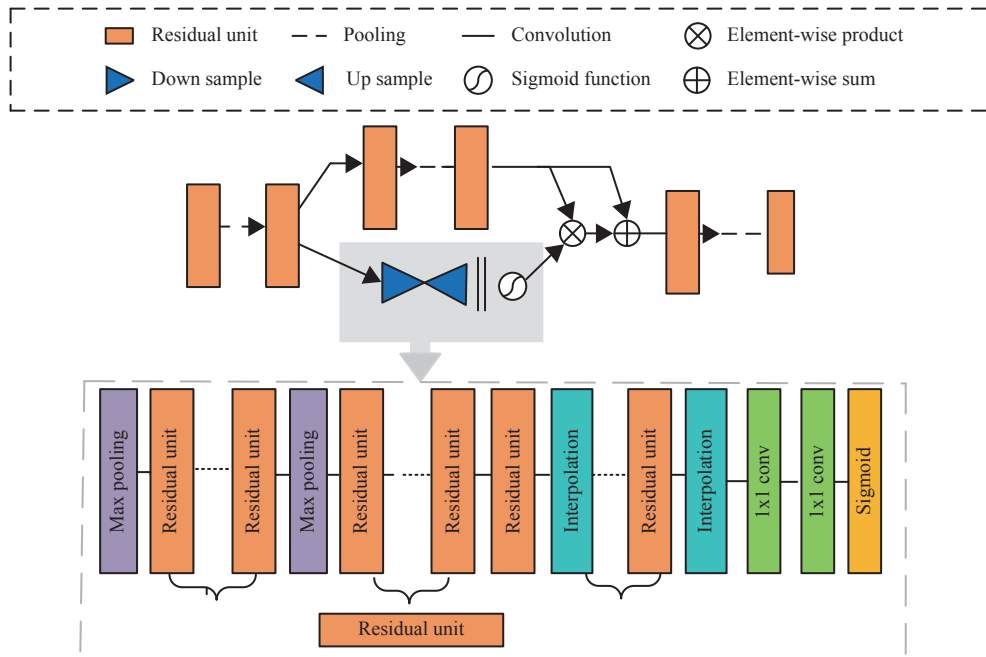


图 2 残差注意力模块

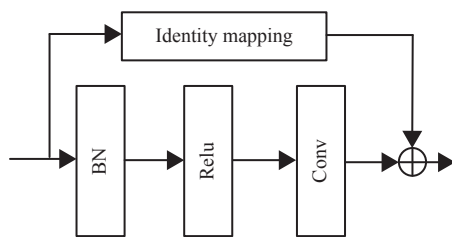


图 3 残差块模型的设计

1.3 U-Net 判别器

Pix2Pix 中的马尔科夫判别器完全由卷积层构成并用于风格迁移的对抗网络中, 需要在内容和纹理两部分进行判别, 但在学习语义、结构和纹理不同任务时容易忘记之前的任务。而 U-Net 使用的跳跃链接可以桥接编码器和解码器两个模块, 输出层的通道包含了不同级别的信息, 因此可以反馈给生成器更多的信息, 同时每个像素的基础上进行分类去判别真实和虚假的全局和局部决策, 提高了判别器的网络性能。

本文使用 U-Net 网络来判别输入图像的真假, 判别器是由编码器和解码器两部分构成的 U-Net 网络。编码器有 5 个大小为 $4 \times 4 \times 4$ 的卷积核, 步幅为 2, 每个卷积后是一个 ReLU 函数, 随后加入 Batch-Norm。解码器先进行上采样, 特征图的大小变为原来的两倍, 然后经过大小为 $4 \times 4 \times 4$, 步幅为 1 的反卷积, 同样加入 ReLU 函数、Dropout 和 Batch-Norm 层。依次通过 4 个同样的反卷积运算后再经

过一次上采样和一个步幅为 2 的卷积, 将图片输出转换为 $2 \times 2 \times 2$ 大小。同时在两个模块之间建立了跳跃连接, 最后得到判别器的输出概率来判断生成的肺结节。

2 实验

2.1 数据集

本文采用 LUNA16 公开数据集^[18], LUNA16 数据集包含 888 个低剂量肺部 CT 影像。其中带有医生标注且直径大于 3 mm 并且相近的结节融合的肺结节共 1 186 个。为了使生成结节的效果更优, 从中挑选直径 1~1.6 cm 的肺结节, 同时为增加训练样本数量, 对数据进行翻转和旋转数据的增强方式。结合考虑机器硬件条件的限制, 最终得到训练集有 5 152 个结节训练样本。

2.2 实验细节

实验环境以 Python3.6 为编程语言, 编码在 keras, Tensorflow 深度学习框架, 采用 GPU 为单张 RTX3090 显卡行。网络优化采用 Adam 优化器, 学习率为 0.000 2, 动量参数分别为 $\beta_1=0.5$, $\beta_2=0.999$; 损失函数中 λ 设置为 10, 批处理大小为 16, 进行 200 轮次生成。

2.3 评价指标

实验采用 FID^[19], PSNR^[20], SSIM^[21] 这 3 种评价指标来评估生成结果。FID 通过计算生成样本与

真实样本特征空间的距离来衡量两幅图像之间的相似度。FID 数值越小说明两者越接近, 意味着生成图片的质量较高、多样性丰富。SSIM 是一种衡量两幅图像结构相似度的指标, 对图像局部变化敏感, 取值范围在 0~1 之间, SSIM 数值越大说明两张图片越相似。PSNR 是用于衡量图像质量的重要指标, PSNR 数值越大说明失真越少, 生成图像质量越好。

2.4 实验结果和分析

2.4.1 掩膜大小的选定

为了选取合适大小的掩膜以达到模型最优效果, 实验分别选取了 $8 \times 8 \times 8$ 、 $12 \times 12 \times 12$ 、 $16 \times 16 \times 16$ 、 $20 \times 20 \times 20$ 及 $24 \times 24 \times 24$ 这 5 种不同尺寸的掩膜。在 Pix2Pix 模型上进行实验, 生成图像如图 4 所示, 对应 FID 评价指标如表 1 所示。图 4 第一行为不同尺寸的掩膜图像, 第二行为生成图像, 第三行为原图像。由图可知, $8 \times 8 \times 8$ 和 $12 \times 12 \times 12$ 尺寸的掩膜过小, 对于较大结节不能完整覆盖, 导致生成结果与原图过于相似, 虽然 FID 值较小, 但其增强图像应用于深度学习训练网络可能会导致过拟合。 $20 \times 20 \times 20$ 和 $24 \times 24 \times 24$ 大小的掩膜过大, 生成结节边界模糊, 对应于表 1 发现掩膜越大其 FID 数值也越大, 因此最终选取 $16 \times 16 \times 16$ 大小的掩膜作为源域图像。

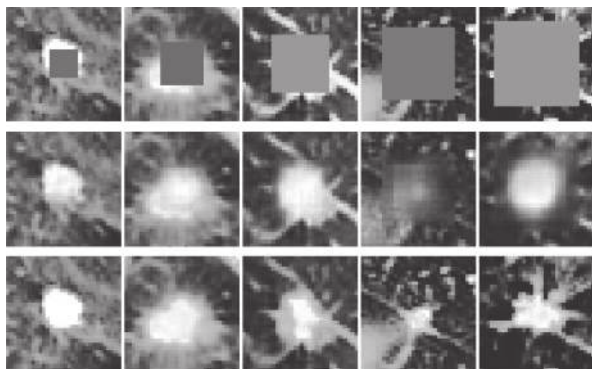


图 4 不同大小掩膜的选取

表 1 不同掩膜尺寸的指标结果

掩膜大小	$8 \times 8 \times 8$	$12 \times 12 \times 12$	$16 \times 16 \times 16$	$20 \times 20 \times 20$	$24 \times 24 \times 24$
FID	80.376 2	83.891 4	86.848 8	88.657 2	93.452 1

2.4.2 消融实验

为了验证残差注意力机制模块, 改进 U-Net 判

别网络的有效性, 进行消融实验逐一验证。本文以 Pix2Pix 为 Baseline, 实验结果如表 2 所示。首先, 在 Baseline 中添加残差注意力模块, 用 U-Net 判别器替换马尔科夫判别器, 改进后的结果由第二、第四组实验可知, 3 个指标均有提升, 这说明在生成器中嵌入残差注意力模块可以更精细地捕捉各类特征信息, U-Net 判别器在学习不同任务的同时协助生成器生成更优的图片。其次, 在第二组实验的基础上重新设计了残差块, 第二、第三组实验结果对比显示, 改进后的模块结果在评价指标上有降低, 但指标变化微小, 反映出设计后的残差块并没有降低网络生成效果, 同时也达到了不同细节特征良好生成的注意力效果, 但与第二组实验相比减轻了网络深度和参数量, 训练时间得到大幅度的优化, 因此本文最终采用设计后的注意力模块。最后一组实验中, 本文算法 RAU-GAN 与 Pix2Pix 相比较, FID 降低 1.861 3, PSNR 提高 0.32, SSIM 提高 0.000 3, 表明了本文模型的有优越性。

表 2 消融实验

组别	方法	Att	Att(res)	U-Net	FID	PSNR	SSIM
第一组	Pix2pix(Baseline)				86.848 8	28.594 4	0.953 2
第二组	Pix2pix+Att	√			85.410 3	28.788 4	0.953 3
第三组	Pix2Pix+Att(res)		√		85.451 2	28.781 5	0.953 3
第四组	Pix2Pix+Unet			√	85.692 8	28.746 9	0.953 4
第五组	RAU-GAN	√		√	84.398 7	28.914 4	0.953 5

2.4.3 在 DeepLesion 数据集上的实验结果

为了验证本文方法的通用性, 在 Deep Lesion^[22] 公开数据集上进行验证。Deep Lesion 包含了多种病变类型的图像, 从中选取 2 209 张带有标注的肺结节图像作为训练样本, 利用关键切片的标注信息裁剪 $32 \times 32 \times 32$ 大小的肺结节图片, 使病灶位置位于切片的中央, 随后进行掩膜处理。将掩膜数据和裁剪后的肺结节作为源域和目标域图片输入, RAU-GAN 模型的生成结果如图 5 所示, 结果表明该模型能够在 Deep Lesion 数据集生成较好的结节影像。将 Pix2Pix 与 RAU-GAN 对该数据集生成结果进行指标评估, 对应结果如表 3 所示, 由于数据集仅 2 209 张, 相较增强后的 LUNA16 数据集少, 因此评价指标较表 2 有稍微降低, 但 RAU-GAN 与 Pix2Pix 相比, 3 个评价指标均明显提升, 因此证明了 RAU-GAN 模型具有泛化能力。

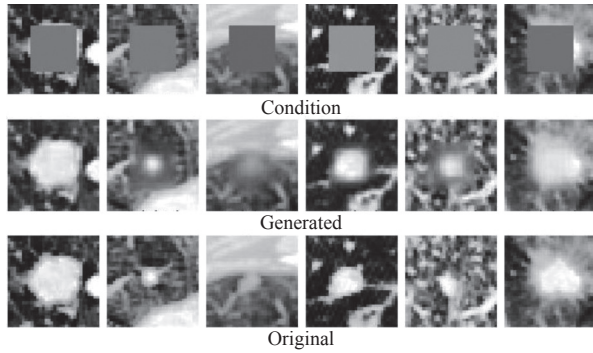


图 5 DeepLesion 数据集上的实验结果

表 3 不同模型在 Deep Lesion 数据集上的结果

方法	FID	PSNR	SSIM
Pix2pix	87.848 8	28.584 4	0.952 9
RAU-GAN	85.548 4	28.703 3	0.953 2

2.4.4 不同模型实验结果及分析

不同模型在 LUNA16 数据集生成肺结节的实验结果如图 6 所示。图 6a 是由 Pix2Pix 模型生成的肺结节图片, 每一列为一组, 整张图片包含 3 组结果。其中第一行是处理后的掩码图片, 第二行图像为生成的肺结节图片, 第三行真实图像为目标域图片。图 6b 是由 CycleGAN 模型生成的肺结节图片, 包含的 6 张图片为一组生成结果, 第一行包含的 3 张图片是目标域图片经由第一个生成器生成的图片, 再经由第二个生成器将生成的图片尽可能地恢复回原始图片, 第二行是由源域图片进行同样的两次生成过程, 最终生成的肺结节图像为第二行第二列。图 6c, 6d, 6e 分别为 U-GAT-IT^[23], DualStyleGAN^[24]和本文提出的 RAU-GAN 模型生成的肺结节图片, 包含的 9 个图片分布情况与图 6a 相同。由图片显示结果可知, CycleGAN 生成的图像中掩码依旧模糊存在, 没有生成结节形状, 因此生成图片无法作为扩充的数据集使用。考虑到 CycleGAN 两个生成器生成图片之间的循环一致损失, 只要求第二次生成结果与原图像越相似越好, 因此生成肺结节的过程并没有明确约束条件, 意味着生成结果可能存在多解, 所以生成肺结节效果较差。其他模型生成结果明显优于 CycleGAN, 但图片显示, Pix2Pix, U-GAT-IT 和 DualStyleGAN 在生成肺结节的边界较模糊, RAU-GAN 生成结果包含更多的细节信息。

表 4 列出了 5 种模型在不同指标上的结果。RAU-GAN 模型在 3 个指标上均优于另外 4 个模型, 验证了本文提出方法的有效性。

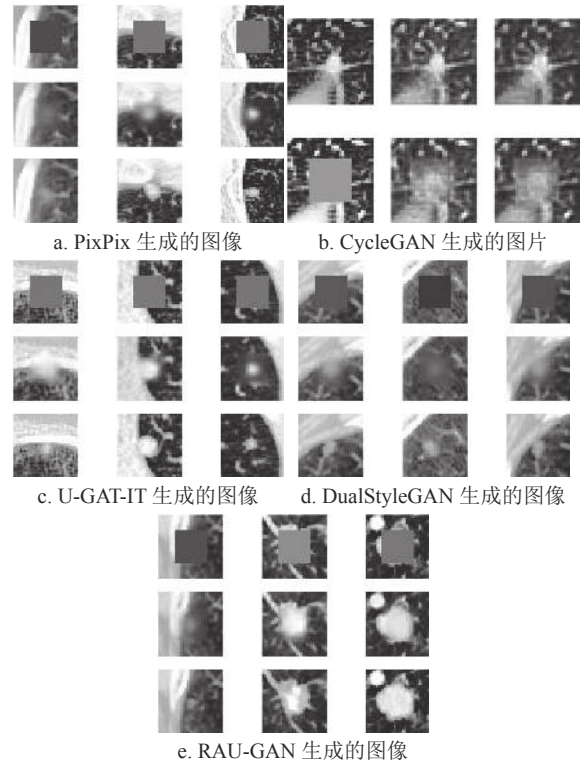


图 6 不同模型生成的肺结节图像

表 4 不同模型在数据集上的指标生成结果

方法	FID	PSNR	SSIM
Pix2Pix	86.848 8	28.594 4	0.953 2
CycleGAN	87.372 8	28.442 5	0.952 1
U-GAT-IT	86.838 5	28.595 2	0.953 2
DualStyleGAN	86.061 5	28.684 7	0.953 3
RAU-GAN	84.987 5	28.914 4	0.953 5

3 结束语

本文针对带标签的肺 CT 数据集匮乏的问题, 提出了基于残差注意力机制和 U-Net 框架进行的生成算法。该模块生成器通过引入残差注意力机制, 堆叠的注意力模块对不同特征信息赋予高的权重, 有效地生成细节信息。此外, 通过对残差块进行重新设计来降低生成网络模型的复杂度, 避免了网络梯度消失的问题。对于判别器网络, 通过对 U-Net 网络结构进行重新设计来进一步提高判别性能。本文使用 FID、PSNR 和 SSIM 作为评价指标, 来保证生成结果的相似性和生成质量, 生成结果的唯一性和差异性也需保证, 以避免后续深度模型训练过拟合的问题, 因此未来可进一步探讨相关的图像评价指标来保证生成图像的真实性和唯一性。此外, 如果不进行 CT 影像裁剪, 生成的结果往往效果不佳, 同时, 实验生成结果表明, 结节越

大生成效果越好,因此如何将当前基于生成式对抗网络的模型更好地扩展到大背景下精细地生成小目标,也是下一步的研究重点。

参 考 文 献

- [1] 尹周一,王梦圆,游伟程,等. 2022 美国癌症统计报告解读及中美癌症流行情况对比[J]. 肿瘤综合治疗电子杂志, 2022, 8(2): 54-63.
YIN Z Y, WANG M Y, YOU W C, et al. Interpretation of the 2022 American Cancer Statistics Report and comparison of cancer prevalence in China and the United States[J]. Electronic Journal of Integrative Oncology Therapy, 2022, 8(2): 54-63.
- [2] YUN S, HAN D, OH S J, et al. Cutmix: Regularization strategy to train strong classifiers with localizable features[EB/OL]. [2022-4-11]. <https://arxiv.org/abs/1905.04899v1>.
- [3] WANG H, HAO W, ZHILI S. Improved mosaic: Algorithms for more complex images[J]. *Journal of Physics:Conference Series*, 2020, 1684(1): 012094.
- [4] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks[J]. *Communications of the ACM*, 2020, 63(11): 139-144.
- [5] LIANG J, YANG X, HUANG Y, et al. Sketch guided and progressive growing GAN for realistic and editable ultrasound image synthesis[J]. *Medical Image Analysis*, 2022, 79: 102461.
- [6] DAR S U, YURT M, KARACAN L, et al. Image synthesis in multi-contrast MRI with conditional generative adversarial networks[J]. *IEEE Transactions on Medical Imaging*, 2019, 38(10): 2375-2388.
- [7] SHI H, LU J, ZHOU Q. A novel data augmentation method using style-based GAN for robust pulmonary nodule segmentation[C]//2020 Chinese Control And Decision Conference (CCDC). Hefei: IEEE, 2020: 2486-2491.
- [8] 李阳,高轶奇. 基于数据增强及注意力机制的肺结节检测系统[J]. 北京邮电大学学报, 2022, 45(4): 25-30.
LI Y, GAO S Q. Lung nodule detection system based on data augmentation and attention mechanism[J]. *Journal of Beijing University of Posts and Telecommunications*, 2022, 45(4): 25-30.
- [9] MIRSKY Y, MAHLER T, SHELEF I, et al. CT-GAN: Malicious tampering of 3D medical imagery using deep learning[C]//28th USENIX Security Symposium. Santa Clara: [s.n.], 2019: 461-478.
- [10] ZHANG H, GOODFELLOW I, METAXAS D, et al. Self-Attention generative adversarial networks[C]// International Conference on Machine Learning. [S.l.]: PMLR, 2019: 7354-7363.
- [11] EMAMI H, ALIABADI M M, Dong M, et al. SPA-GAN: Spatial attention gan for image-to-image translation[J]. *IEEE Transactions on Multimedia*, 2020, 23: 391-401.
- [12] WENG W, ZHU X. INet: Convolutional networks for biomedical image segmentation[J]. *IEEE Access*, 2021, PP(99): 16591-16603.
- [13] SCHNFELD E, SCHIELE B, KHOREVA A. A U-Net based discriminator for generative adversarial networks[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 8204-8213.
- [14] ZHAO W, ZHAO Y, FENG L, et al. Attention enhanced serial U-net++ network for removing unevenly distributed haze[J]. *Electronics*, 2021, 10(22): 2868.
- [15] GUI J, SUN Z, WEN Y, et al. A Review on generative adversarial networks: Algorithms, theory, and applications[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2021, 35(4): 3313-3332.
- [16] ISOLA P, ZHU J Y, ZHOU T, et al. Image-to-image translation with conditional adversarial networks[C]//IEEE Conference on Computer Vision & Pattern Recognition. Honolulu: IEEE, 2017: 5967-5976.
- [17] FEI W, JIANG M, CHEN Q, et al. Residual attention network for image classification[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE, 2017: 6450-6458.
- [18] SETIO A, TRAVERSO A, De Bel T, et al. Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the LUNA16 challenge[J]. *Medical Image Analysis*, 2016, 42: 1-13.
- [19] HEUSEL M, RAMSAUER H, UNTERTHINER T, et al. GANs Trained by a two time-scale update rule converge to a local nash equilibrium[EB/OL]. [2022-5-10]. <https://arxiv.org/pdf/1706.08500.pdf>.
- [20] TONG Y B, ZHANG Q S, QI Y P. Image quality assessing by combining PSNR with SSIM[J]. *Journal of Image and Graphics*, 2006, 11(2): 1758-1763.
- [21] ZHOU W, BOVIK A C, SHEIKH H R, et al. Image quality assessment: From error visibility to structural similarity[J]. *IEEE Trans Image Process*, 2004, 13(4): 600-612.
- [22] YAN K, WANG X, LU L, et al. Deep Lesion: Automated mining of large-scale lesion annotations and universal lesion detection with deep learning[J]. *Journal of Medical Imaging*, 2018, 5(3): 036501.
- [23] KIM J, KIM M, KANG H, et al. U-gat-it: Unsupervised generative attentional networks with adaptive layer-instance normalization for image-to-image translation[EB/OL]. [2022-5-12]. <https://arxiv.org/abs/1907.10830v2>.
- [24] YANG S, JIANG L, LIU Z, et al. Pastiche master: Exemplar-based high-resolution portrait style transfer[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2022: 7693-7702.

编辑 叶 芳