

多模态弱监督学习在肝癌图像生成与分割中的应用

潘依乐, 高永彬

上海工程技术大学电子电气工程学院, 上海 201620

【摘要】针对MRI软组织对比度高但不是胸部成像护理标准,导致难以获得足够多专家标注的MRI数据的问题,通常将CT转换成MRI图像。由于难以获取对应模态的CT和MRI图像,结合生成式对抗网络的结构特点,提出CSCGAN生成网络模型。该模型以CycleGAN作为框架,由于CycleGAN可能存在模式坍塌问题,同时StyleGAN2能够控制合成图像的样式和特征细节,实现高分辨率图像的合成,因此将其融入到CycleGAN中,重构了网络的生成器。同时为了减少外部干扰,引入了噪声模块,另外为了防止肿瘤在转换时丢失,修改了网络的鉴别器结构,并加入了混合注意力机制。实验结果显示,与文中其他方法相比,该模型生成的样本图像在Dice相似系数、Hausdorff距离、体积比和平均交并比各项指标上均有所提升,该方法有效实现了肝脏肿瘤病变图像的模态转换,生成的数据能够提高分割网络的准确性。

【关键词】医学图像处理;模态转换;肝脏肿瘤分割;生成式对抗网络;混合注意力机制

【中图分类号】R318;TP391

【文献标志码】A

【文章编号】1005-202X(2024)01-0008-10

Application of multimodal weakly-supervised learning in image synthesis and segmentation of liver cancer

PAN Yile, GAO Yongbin

School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China

Abstract: Although it has high resolution for soft tissues, magnetic resonance imaging (MRI) is not the standard for chest imaging, which results in an insufficient amount of expert-annotated MRI data. Therefore, CT image is usually converted into MRI image. To overcome the difficulty of obtaining the corresponding modal CT and MRI images, a CSCGAN model with CycleGAN as the framework is proposed based on the structural characteristics of generative adversarial networks. Considering the possibility of mode collapse in CycleGAN, StyleGan2 which can control the style and feature details of the synthetic image and realize the synthesis of high-resolution images is integrated into CycleGAN for reconstructing the generator. A noise module is introduced to reduce external interference. In addition, in order to prevent the loss of tumors during conversion, the discriminator structure of the network is modified, and a mixed attention mechanism is added. Experimental results show that compared with the images generated by other methods, those generated by the proposed model are improved in Dice similarity coefficient, Hausdorff distance, volume ratio and mean intersection over union, indicating that the proposed method can effectively realize the mode conversion of liver tumor images, and that the generated data can improve the segmentation accuracy.

Keywords: medical image processing; mode conversion; liver tumor segmentation; generative adversarial network; mixed attention mechanism

前言

【收稿日期】2023-07-09

【基金项目】国家工业和信息化部 and 卫健委 5G 医疗示范项目;广州市科技项目(202206010093);广东省重点研发项目(2020B010165004)

【作者简介】潘依乐, 硕士研究生, 研究方向: 电子信息, E-mail: panyile@126.com

【通信作者】高永彬, 博士, 副教授, 硕士生导师, 研究方向: 视觉SLAM技术、三维视觉分析、自然语言处理、知识图谱、无人机/无人车、智能安防、智慧医疗、智能制造, E-mail: gaoyongbin@sues.edu.cn

基于磁共振的放射治疗要求精确的肿瘤分割。由于磁共振成像(Magnetic Resonance Imaging, MRI)不是胸部成像的标准,因此很难获得足够多的具有专家标注的MRI图像集来训练深度学习方法。计算机断层扫描(Computed Tomography, CT)和MRI多模态转换是人工智能技术应用于医学影像最有前途的领域之一,CT和MRI图像转换可以表述为从原始CT图像到目标MRI图像的映射问题,MRI和CT之间的交叉通道图像配准是一个具有挑战性的问题,因为不同的成像机制会导致组织或器官外观的高度变异性,缺乏一个通用的规

则来比较这类图像。CT和MRI图像中的像素强度、体素大小、图像方向和视场也不同,导致多模态配准效果不如单模态配准效果好。在未配对图像数据生成中,循环一致性生成对抗网络(Cycle-Consistent Generative Adversarial Networks, CycleGAN)可以获得较好的结果^[1],但仍有提升的空间。NVIDIA提出的StyleGAN2用于生成高质量的真实图片^[2],基于此,本文结合CycleGAN和StyleGAN2模型的特点,提出一种基于CSCGAN的未配对多模态图像转换方法。其主要思想是将StyleGAN2精细的控制特点融合到CycleGAN的生成器网络中,构造新的生成器和鉴别器,同时添加混合注意力机制。

1 研究现状

随着MRI引导放射治疗(MRIgRT)的发展,由MRI扫描仪集成的加速器优势明显,逐渐取代了利用CT进行的影像引导放射治疗技术(Image-Guided Radiotherapy Technique, IGRT)。MRIgRT有几个明显的优势,首先,MRI图像中软组织的对比度明显优于CT图像,可以提供更清晰的肿瘤病理和周围区域的细节^[3-4]。其次,MRI无需额外放疗,可以实现实时成像,可以基于软组织可视化进行快速自适应规划和约束投放控制,使用精确的MRI指导放疗能够进一步降低辐射剂量。

医疗图像常见的处理需求包括检测、分割识别医学图像中的区域与对象等^[5]。而基于磁共振的放射治疗要求精确的肿瘤分割。不同模式之间的标注可能会对基于软组织的IGRT标注产生不利影响。对于CT转换MRI,已经有许多合成方法被提出,如Emami等^[6]提出一种以残差网络作为生成器的模型,在数秒内实现MRI图像生成高质量CT图像。Emami等^[7]在生成器中加入图像域翻转平移的全局流和保留两域之间不一致区域的局部流,更好地保留了生成图像的细节信息。Qi等^[8]和Lei等^[9]从体系结构的角度对CycleGAN进行了修改。Yang等^[10]和Ge等^[11]从损失函数的角度对CycleGAN进行了修改。

对于MRI-CT配准,与CT-MRI配准相比,Zachiu等^[12]展示了更好的单模配准结果^[13]。为了解决具有挑战性的多模态图像配准问题,已经提出了多种方法来使用成对的MRI和CT数据进行MRI到CT图像的转换。这些模型包括高斯混合回归模型、随机森林回归模型、基于分割的方法。基于地图集的方法和基于学习的方法,如Xiang等^[14]设计了一个深度嵌入卷积神经网络(Convolutional Neural Network, CNN)来学习MRI到CT映射等。另外,还提出了一些基于非监督深度学习的方法用于MRI-CT合

成^[15-21]。如Nie等^[22]在pix2pix网络中添加了基于梯度的损失函数,对图像边缘进行了强化,实现了上下文感知,可用于MRI到CT图像合成。然而,CT图像中丰富的解剖信息的丢失使得精确的配准变得困难,因为足够的解剖细节也可以引导精确图像配准^[23]。Wolterink等^[24]提出了将GAN应用在低剂量CT到常规多种剂量的转换。Jiang等^[25]在MRI与CT的相互转换中引入一种新的靶标特异性肿瘤感知损失算法,来加强对容易丢失的肿瘤结构的学习,使得肿瘤信息在图像的转换过程中能得到更好的保留。

以上几种方法大部分使用CT和MRI相互对应的数据作为训练数据,但在实际应用中,很难收集到对应的CT和MRI数据,为了解决不配对医学图像的跨模态转换,Dar等^[26]提出了一种多对比度的MRI图像合成端到端框架,通过Patch-GAN和CycleGAN来实现,但是该网络合成的是三维MRI图像。Chartsias等^[27]使用CycleGAN模型和未配对的心脏数据实现不同患者的CT图像生成MRI图像。然而,这项工作不能直接定量评价成由未配对的心脏图像生成的MRI图像,因为很难从一直跳动的心脏中获得MRI和CT的成对图像。Jin等^[28]通过使用配对和非配对的脑CT/MRI数据生成MRI图像,提高了放疗规划中基于CT目标勾画的准确性,其采用的是一种非监督学习神经网络,但并没有比较传统的监督学习方法和非监督学习方法在MRI图像生成任务中的效果差异。Abu-Srhan等^[29]基于Wasserstein GAN(WGAN)模型提出一种无监督注意引导的GAN模型,实现MRI图像与CT图像的双向转换。

2 研究方法

本文提出了一种新型生成对抗网络模型CSCGAN,该网络利用CycleGAN作为框架,添加精细控制结构形成新的生成器,同时修改鉴别器结构,并将生成数据和真实数据按照不同结合比例放入不同的分割网络进行分割,分析生成数据的质量以及在数据扩充方面对于分割网络的增益作用。本文模型总体框架流程图如图1所示。

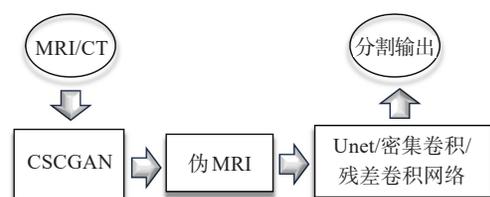


图1 生成和分割模型框架流程图

Figure 1 Framework flowchart of image synthesis and segmentation model

本研究的生成与分割总体步骤如下:(1)借鉴CycleGAN不需要配对数据和StyleGAN2能够精细控制生成细节的特点,提出了CSCGAN模型,用于生成接近真实图片的数据。(2)使用Chaos医学数据集来预训练分割模型,预先获得合适的预训练神经网络模型参数,从而提高分割准确性。(3)将生成模型生成的数据放入分割网络。分割网络损失函数为Dice loss系数,该损失函数在医学图像分割数据不均衡问题上可以有效降低影响,从而改善分割模型的性能。(4)通过不同策略验证生成模型,分析各项结果指标。肝脏肿瘤图像生成与分割模型结构如图2所示。

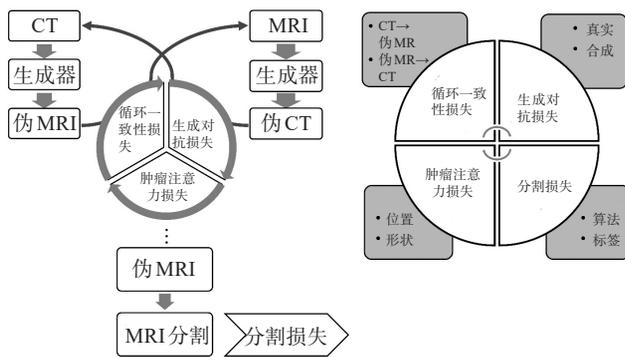


图2 肝脏肿瘤图像生成、分割模型示意图
Figure 2 Diagrams of liver tumor image synthesis and segmentation model

2.1 生成式对抗网络模型

作为一种深度学习模型,生成式对抗网络经过两个生成网络和判别网络模块的互相博弈,获得优异的输出结果。生成网络的输入是随机的噪声,由噪声生成图片,输出要求是最大化接近真实的训练数据,其作用是尽量生成接近真实的图片去隐瞒判别网络。判别网络以真实的数据或者生成网络输出结果作为输入,其作用是辨别输入的是真实的数据还是生成的假数据。两个网络相互对抗,不断调整参数。

2.1.1 CycleGAN 模型 CycleGAN是一个由两对生成器及判别器组成的环形结构,不需要完全配对的数据。在生成器中,分为编码和解码两个部分。如图3所示, X 表示 X 输入域的图像, Y 表示 Y 目标域的图像。生成器 G 表示 X 到 Y 的变换,生成器 F 表示 Y 到 X 的变换和逆变换, D_X 和 D_Y 用来判别 F 和 G 的性能。 X 通过 G 生成 Y ,再通过 F 重新构造出 X 原图像;同等原理, Y 最终重新构造出 Y 原图像。将原域图片 X 送入($X \rightarrow Y$ 方向) G 生成目标域图片 \hat{Y} , D_Y 判别 \hat{Y} 的真假,然后再将生成的目标域图片 \hat{Y} 送入($Y \rightarrow X$ 方

向) F ,逆变换生成原域图片 \hat{X} ,对真实 X 和生成 \hat{X} 进行判断。反之,反向生成器($Y \rightarrow X$) F 通过判别器 D_X 进行博弈。

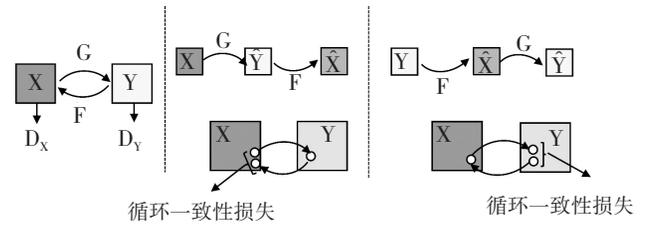


图3 CycleGAN结构
Figure 3 Structure of CycleGAN

2.1.2 StyleGAN2 模型 StyleGAN是一种新的生成网络,由Karras等^[30]提出。可以通过无监督式的自动学习对图像的高层语义属性做一定的解耦分离,也可以直观地按照特定尺度做到一定程度上的控制合成。但是,StyleGAN在生成图像的过程中有时包含斑点似的伪影(artifacts),而StyleGAN2重点修复artifacts问题,并进一步提高了生成图像的质量和训练的速度。StyleGAN2与StyleGAN在结构上的区别如图4所示。

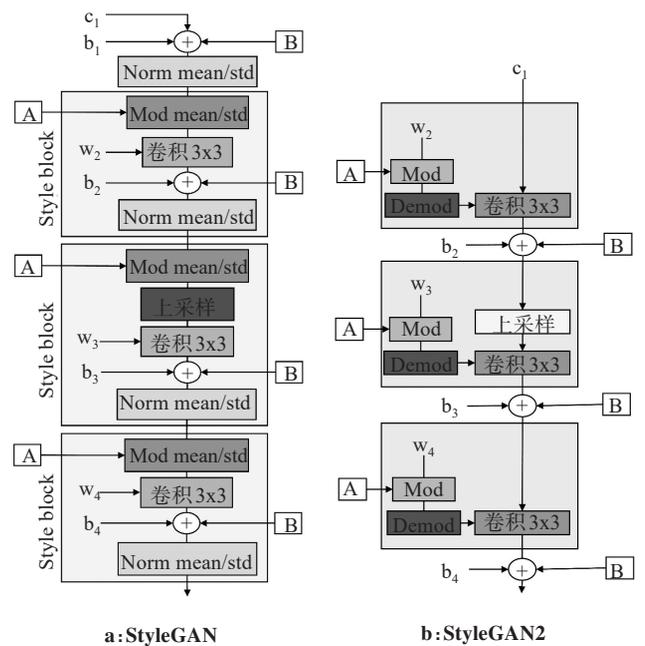


图4 StyleGAN与StyleGAN2的结构对比
Figure 4 Structure comparison between StyleGAN and StyleGAN2

2.2 CSCGAN 网络模型

神经网络的训练需要大量的数据来支撑,由于医疗图像数据的收集难度大,需要专家进行手动分

割,耗时耗力,且分割标准受多种因素影响,因此,本研究采用了生成对抗网络生成肝脏肿瘤数据,扩充数据集,提高分割网络的分割精度。

由于配对的CT和MRI图像难以获取,故以CycleGAN作为框架,学习CT到MRI的转换,但是由于CycleGAN没有对图像建立明确的模型,可能存在模式坍塌(mode collapse)问题,导致生成的图像质量不高。而StyleGAN2能够控制合成图像的样式和特征的精细,实现高分辨率图像的合成,因此借鉴了

StyleGAN2的生成器精细控制特点,将其融入到CycleGAN中,提高生成样本的精度。加入噪声模块,结合合成网络学习到的样式细节与接收得到的噪声,得到新的肝脏肿瘤数据,使数据集的随机性和多样性得以扩大。另外,为了防止大量变异的结构(例如肿瘤)在CT到MRI图像转换时丢失,修改了鉴别器结构,在其中加入了混合注意力机制。图5显示了CSCGAN的网络结构。

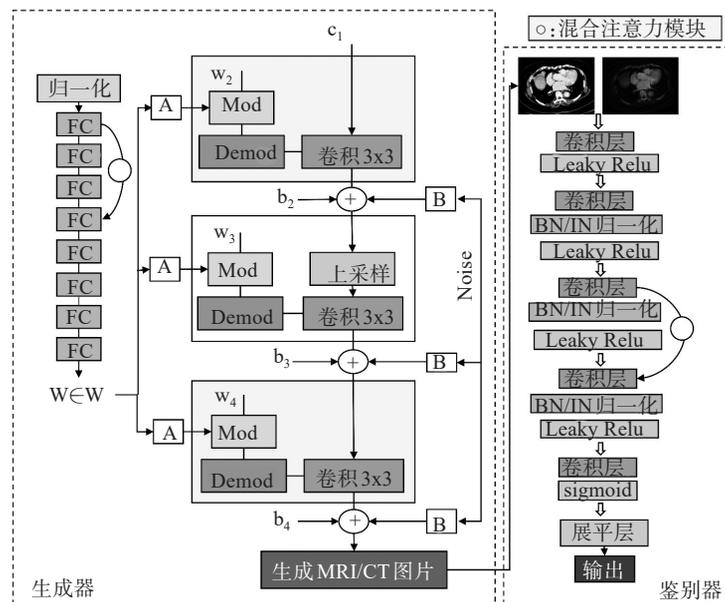


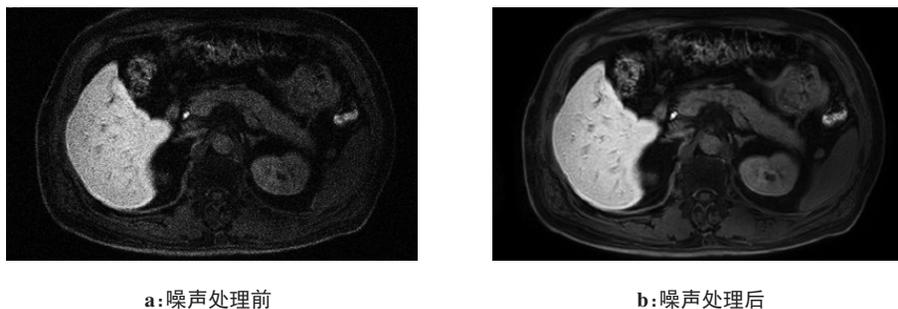
图5 CSCGAN网络结构图

Figure 5 Structure of CSCGAN

研究模型CSCGAN借鉴了CycleGAN的处理模式,对网络的生成器进行了修改,融入了StyleGAN2的样式控制模式,并修改了噪声模块和鉴别器的结构。主要的修改内容有以下几点。

(1)使用原始的噪声结构,在生成的图像中会得到许多额外的噪点,为了避免其不利的影响,将噪声模块

移到style block内,噪声模块仅加在第一层。如图6所示,改进后生成的图像去掉了绝大部分噪点,获得了更优秀的生成数据。由于不同时期、不同类别的肝癌较为相似,病变边缘不清晰等特殊特性,对于数据的准确性要求非常高,本文的改进噪声部分可为肝脏临床区分病变区域提供更精确的样本数据。



a: 噪声处理前

b: 噪声处理后

图6 使用原始噪声模块和改进后噪声模块结果对比图

Figure 6 Comparison before and after noise removal

(2)修改了CycleGAN鉴别器的结构。每个卷积层的卷积核数量分别是128、256、512、1024。为了稳定鉴别器网络训练,除首层和最后层外,在其余卷积

层中使用leaky ReLU函数而不是原始的ReLU函数。鉴别器结构如图7所示。

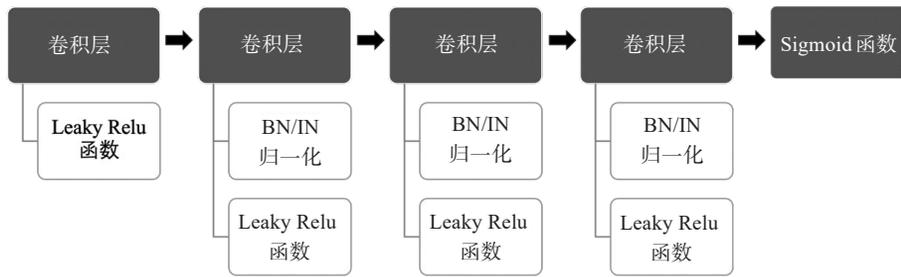


图7 鉴别器结构

Figure 7 Discriminator structure

(3)在鉴别器结构中添加混合注意力模块。由于数据的分辨率高,容易产生不稳定的训练,导致生成的数据受到影响。为了防止肿瘤在从CT到MRI图像转换时不能完全精准保留,将混合注意力模块融合到特征提取和融合中。本文融合了混合注意力机制,由通道注意力模块为不同通道分配权重,获取CT和MRI图像和其中的显著特征;再由空间注意力机制根据不同区域特征的空间关系为各区域分配权重,并根据图像中权重大的区域及其相关度高的区域进行图像转换。混合注意力公式如下:

$$F' = M_c(F) \otimes F$$

$$F'' = M_s(F') \otimes F' \tag{1}$$

在使用上述方法对数据进行模态转换之后,为了验证转换效果对分割网络的影响,预训练了3种常见的分割网络。

(1)U-Net网络。U-Net由一系列卷积块组成,每个卷积块由卷积、批处理归一化和ReLU激活函数组成。跳过连接用于连接高级和低级特征。最大池化层和向上池化层用于下采样和上采样来调整要素图的分辨率大小。在本研究采用的U-Net网络结构中,使用了4个最大池化层和4个向上池化层,U-Net结构如图8所示。

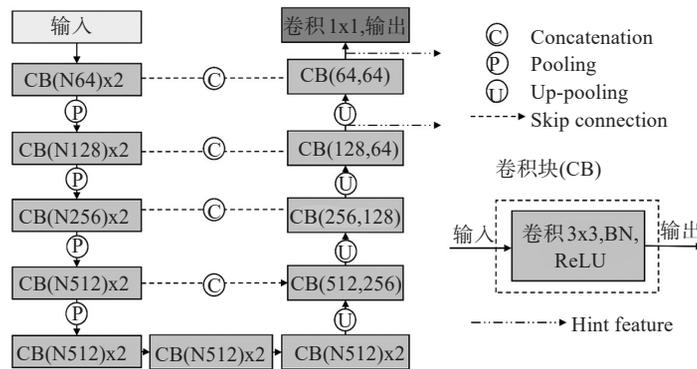


图8 U-Net结构图

Figure 8 Structure of U-Net

(2)Dense-FCN。Dense-FCN由密集块(DB)组成,密集块连续连接从先前层计算的特征图,从而增加特征图的大小。每一层由批量归一化、ReLU和3×3卷积运算组成。向下过渡(Top-Down Transition, TD)和向上过渡(Bottom-Up Transition, TU)分别用于对特征尺寸进行向下采样和向上采样。文中使用

Dense103层结构,使用具有4、5、7、10和12层的密集块进行特征连接,使用5个TD进行特征下采样,使用5个TU进行特征上采样,Dense-FCN结构如图9所示,结果得到9.3 M的参数大小。最后一个大小为256的特征通过1×1卷积后接一个sigmoid函数转换为大小为1的特征图。

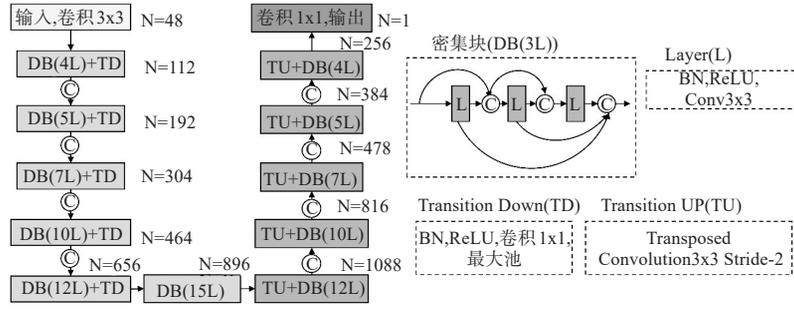


图9 Dense-FCN结构图

Figure 9 Structure of Dense-FCN

(3) 残差连接全卷积网络 (Residual Fully Convolutional Network, RFCN)。为了更快的训练, 使用了FCN-8s结构。网络结构图如图10所示。输入图像首先由 7×7 个卷积滤波器处理, 然后再由 3×3 个卷积滤波器处理。通过最大池进行子采样, 可以减少特征的数量。这些特征随后由带残差块的卷积阵列处理, 被累加并传递到使用sigmoid激活的最终输出层。

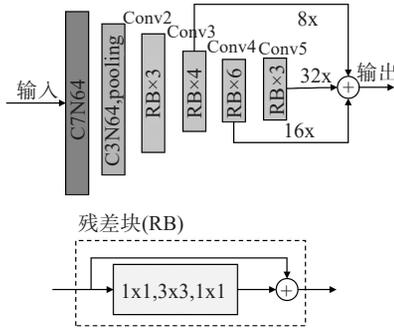


图10 RFCN结构图

Figure 10 Structure of RFCN

2.3 损失函数

(1) 肿瘤注意力损失。肿瘤注意力损失由肿瘤形状损失以及肿瘤位置损失组成, 肿瘤形状损失来自最终网络层的特征图激活 (通过欧几里德距离) 差异最小化, 而肿瘤位置损失迫使两个网络产生相同的肿瘤分割图像 (通过Dice相似系数)^[31]。用公式可以表示为:

$$L_{\text{tumor}}^{\text{shape}} = \frac{1}{C \times W \times H} \left\| \mathcal{O}_{\text{CT}}(x_c) - \mathcal{O}_{\text{CT}}(G_{\text{CT} \rightarrow \text{MRI}}(x_c)) \right\|^2 \quad (2)$$

$$L_{\text{tumor}}^{\text{loc}} = E_{x_c \sim X_{\text{CT}}} \left[y_c \mid G_{\text{CT} \rightarrow \text{MRI}}(x_c) \right] + E_{x_c \sim X_{\text{CT}}} \left[y_c \mid x_c \right] \quad (3)$$

其中, C , W 和 H 是特征通道、宽度和高度, $L_{\text{tumor}}^{\text{shape}}$ 表示肿瘤形状损失, $L_{\text{tumor}}^{\text{loc}}$ 表示肿瘤位置损失。

(2) 生成对抗损失函数。生成器训练以生成接

近真实数据的图像, 鉴别器训练以正确区分真实图像和合成图像, 以最小最大方式协同工作, 这种损失称为对抗损失, 计算为生成伪MRI和伪CT的损失之和: $L_{\text{adv}} = L_{\text{adv}}^{\text{MRI}} + L_{\text{adv}}^{\text{CT}}$ 。用公式表示为:

$$L_{\text{adv}}^{\text{MRI}} = E_{x_m \sim X_{\text{MRI}}} \left[\log \left(D_{\text{MRI}}(x_m) \right) \right] + E_{x_c \sim X_{\text{CT}}} \left[\log \left(1 - D_{\text{MRI}}(G_{\text{CT} \rightarrow \text{MRI}}(x_c)) \right) \right] \quad (4)$$

$$L_{\text{adv}}^{\text{CT}} = E_{x_c \sim X_{\text{CT}}} \left[\log \left(D_{\text{CT}}(x_c) \right) \right] + E_{x_m \sim X_{\text{MRI}}} \left[\log \left(1 - D_{\text{CT}}(G_{\text{MRI} \rightarrow \text{CT}}(x_m)) \right) \right] \quad (5)$$

其中, $x_m \sim X_{\text{MRI}}$ 表示 x_m 来自于真实数据分布 X_{MRI} , $x_c \sim X_{\text{CT}}$ 表示 x_c 来自于真实数据分布 X_{CT} , E 表示对所有数取均值, $\log(D_{\text{MRI}}(x_m))$ 和 $\log(D_{\text{CT}}(x_c))$ 为判别器将真实数据判定为真实数据的概率, $\log(1 - D_{\text{MRI}}(G_{\text{CT} \rightarrow \text{MRI}}(x_c)))$ 和 $\log(1 - D_{\text{CT}}(G_{\text{MRI} \rightarrow \text{CT}}(x_m)))$ 为判别器将生成器生成的虚假数据判定为真实数据的对立面即将虚假数据仍判定为虚假数据的概率。

(3) 循环一致性损失函数。循环一致性损失用来降低对用于训练的原始数据 (CT) 和目标数据 (MRI) 不同模态之间完全对准的要求, 从而允许无监督的跨模态适应学习^[32]。 L_{cyc} 强制网络保留生成图像和原始图像模态之间的空间对应关系, $L_{\text{cyc}} = L_{\text{cyc}}^{\text{MRI}} + L_{\text{cyc}}^{\text{CT}}$, 其中 $L_{\text{cyc}}^{\text{MRI}}$ 和 $L_{\text{cyc}}^{\text{CT}}$ 分别用于合成伪MRI和伪CT图像, 如式(6)所示:

$$L_{\text{cyc}} = E_{x_c \sim X_{\text{CT}}} \left[\left\| (G_{\text{CT} \rightarrow \text{MRI}}(x'_c)) - x_c \right\| \right] + E_{x_m \sim X_{\text{MRI}}} \left[\left\| (G_{\text{MRI} \rightarrow \text{CT}}(x'_m)) - x_m \right\| \right] \quad (6)$$

其中, x'_m 和 x'_c 是从生成器 $G_{\text{MRI} \rightarrow \text{CT}}$ 和 $G_{\text{CT} \rightarrow \text{MRI}}$ 生成的伪CT和伪MRI图像, 伪CT图像只是生成伪MRI图像的附加产物, 生成伪CT图像是为了在合成伪MRI图像时增加额外的正则化。

(4) 研究使用的损失函数计算如下:

$$L_{\text{total}} = L_{\text{adv}} + \lambda_{\text{cyc}} L_{\text{cyc}} + \lambda_{\text{shape}} L_{\text{tumor}}^{\text{shape}} + L_{\text{loc}} L_{\text{tumor}}^{\text{loc}} \quad (7)$$

其中, λ_{cyc} 、 λ_{shape} 、 λ_{loc} 表示每次损失的加权系数。计算的损耗被反向传播以更新网络的权重, 从而从两个网络产生伪MRI和伪CT。

3 实验与结果

3.1 数据集

本文实验所采用的数据集是由广东珠江医院提供且由专业医生进行分割的肝脏肿瘤图像数据集。首先使用3DVWorks专业医疗软件对数据进行标注, 然后使用split软件对数据进行切片处理, 再将切片后的数据转化成png格式的图片以供训练。最终得到由8200张CT图片和5320张MRI图片组成的训练数据, 测试集由3500张CT图片和2280张MRI图片组成。实验过程中, 以训练数据中20%的图片作为验证集, 用来调节模型, 训练集为剩下的80%的图片。

3.2 评估指标

(1) K-L散度。伪MRI与真实MRI图像的相似性使用库尔贝-莱布勒散度(K-L散度)进行评估, K-L散度量化了特征结构(即肿瘤)内伪MRI和真实MRI图像之间整体强度变化的相似性。K-L散度通过以下公式进行计算:

$$D_{KL}(P_{pMRI}|rMRI) = \sum P_{pMRI} \ln \frac{P_{pMRI}}{Q_{rMRI}} \quad (8)$$

其中, P_{pMRI} 和 Q_{rMRI} 表示伪MRI和真实MRI图像中的肿瘤分布, 并且总和是在固定数量的离散强度水平上计算的($N=1000$)。

(2) Dice相似系数(Dice Similarity Coefficient, DSC)和Hausdorff距离。使用DSC和Hausdorff距离计算分割精度。DSC计算公式如下:

$$DSC = \frac{2TP}{FP + 2TP + FN} \quad (9)$$

其中, TP表示真阳性个数, FP表示假阳性个数, FN表示假阴性个数。

Hausdorff距离定义为:

$$H(P, T) = \max \left[\sup_{p \in S_p} \inf_{t \in S_T} d(p, t), \sup_{t \in S_T} \inf_{p \in S_p} d(t, p) \right] \quad (10)$$

其中, P 和 T 分别是专家标注和图像分割体积, p 和 t 分别是 P 和 T 上的点。 S_p 和 S_T 分别对应于 P 和 T 的曲面。另外, 为了消除评估过程中噪声的影响, 采用了95% Hausdorff距离(HD95)。同时计算了相对体积比, 计算方式为:

$$VR = \frac{|V_{as} - V_{gs}|}{V_{gs}} \quad (11)$$

其中, V_{as} 为算法计算的体积, 而 V_{gs} 为人工分割计算的体积。

3.3 实验及结果

3.3.1 使用不同生成网络生成MRI效果 本文使用训练集中肝癌图像来生成更多的样本, 并将其纳入实验中。通过CSCGAN对肝癌病变图像生成2000张伪MRI图片。尽管CSCGAN可以实现无需配对情况下的训练, 但是为了在同等环境下训练, 提前控制变量, 对CSCGAN模型重复训练3次。另外, 在模型训练过程中, 使用 L_2 范数作为损失函数替换原始的不稳定的对数函数, 在初始化过程中使用同样的参数。

图11显示了分别使用GAN、StarGAN、CycleGAN、StyleGAN和本文模型(CSCGAN)生成的肝脏肿瘤MRI图像与专家分割的肝脏肿瘤CT图像的对比图。可以看出, 本文提出的转换方法在伪MRI图像上能够将肿瘤最完整地保存下来。

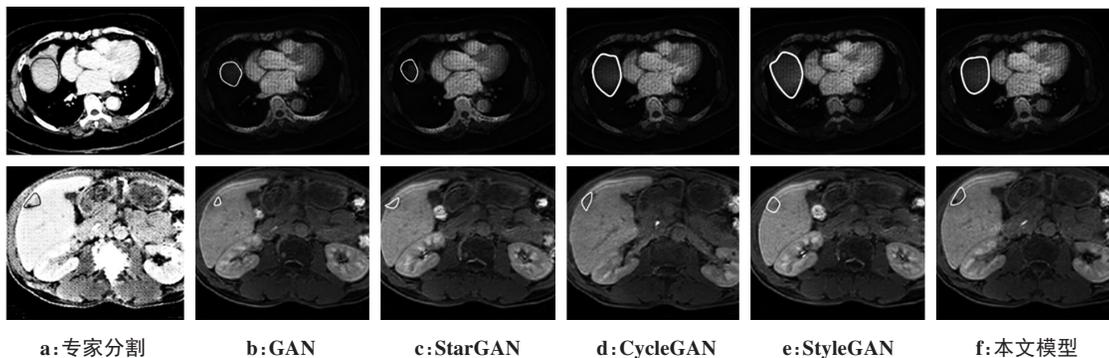


图11 使用不同生成网络生成MRI效果图

Figure 11 Synthetic MRI generated using different networks

本文方法最接近真实MRI成像的单一强度分布。算法生成的肿瘤分布是在CT产生的550个伪MRI图像

的肿瘤内计算的, 而真实MRI分布是从专家分割的MRI图像中使用无效($N=36$)的肿瘤区域获得的。通过足够

大的数据来评估离散强度分布,与GAN(0.52)、StarGAN(0.35)、StyleGAN(0.42)和CycleGAN(1.58)相比,本文方法在伪MRI和真实MRI之间产生的K-L散度最低,为0.087,如图12所示。

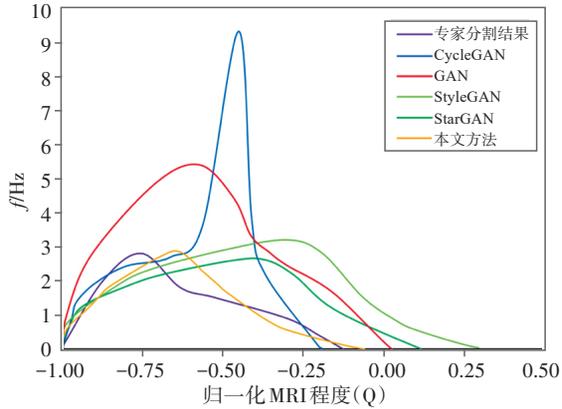


图12 不同方法K-L散度对比图

Figure 12 Comparison of K-L divergence among different methods

3.3.2 生成的MRI数据扩充对基于U-Net、DFCN、RFCN网络的肿瘤分割效果 通过本文研究方法生成的伪MRI数据集与专家分割真实MRI图像相结合,来训练MRI肿瘤分割网络,以评估本文方法的数据扩充策略能否提高多个网络的性能。

图13显示了3个不同患者的分割结果,其中图13a使用随机森林分类器(RF+fCRF)进行训练;U-Net训练仅有的少量真实MRI图像(tMRI)(b);U-Net训练真实MRI和伪MRI图像(t&p MRI),对应于弱监督训练,使用CycleGAN(c)、StyleGAN(d)、GAN(e)和StarGAN(f)产生的交叉模态增强进行训练;图13g显示了仅使用本文方法产生的伪MRI的训练结果,以便在仅使用增强的伪MRI数据时评估该方法的性能,该模式设置对应于无监督域自适应;图13h表示使用本文方法训练真实MRI和伪MRI图像。如图13g、h所示,无监督和弱监督模式方法生成的轮廓(白色)非常接近专家分割(黑色)。

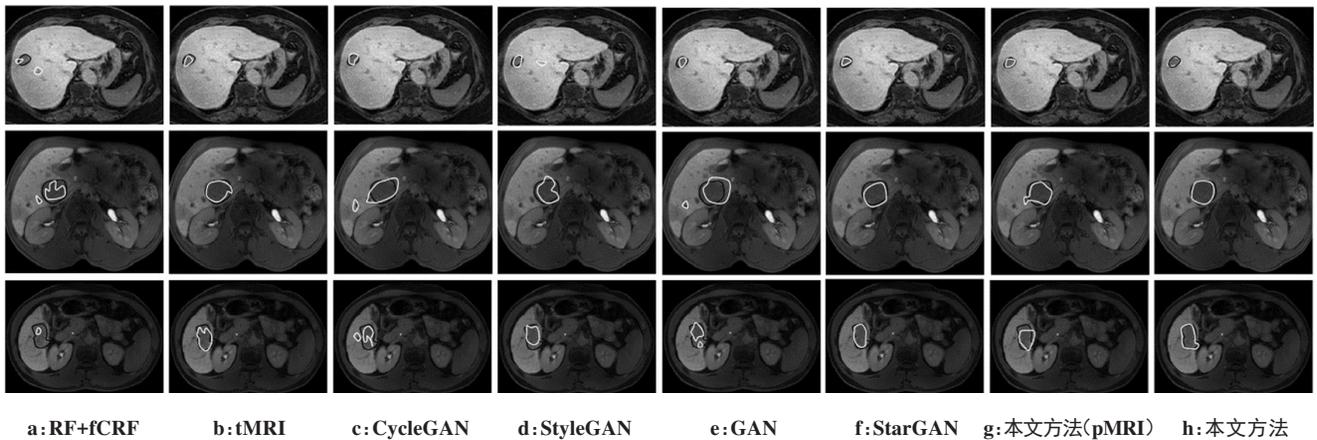


图13 不同生成网络生成数据的U-Net分割结果

Figure 13 U-Net segmentation results using data generated by different networks

使用DSC、Hausdorff距离和VR评估以上几个方法的总体分割精度,结果如表1所示。本文方法(t&p MRI)获得了最高的DSC和最低的HD95以及VR。

另外,评估了两个附加分割网络RFCN和DFCN的分割精度,以评估使用本文方法的数据扩充是否提高了这些网络的分割精度。结果如表2所示,本文方法(t&p MRI)获得最高DSC和最低的HD95以及VR。

为了验证研究模型生成的数据在真实数据不充分时对分割网络分割精度的增益效果,将生成的数据按比例融入到真实数据中。为了控制变量,随机选取10%、20%、50%、100%的真实MRI数据,同时添加生成MRI数据,组成不同比例的数据集,将其归一化后输入到不同分割网络中。

分割结果如表3所示。结果显示,性能随着更充分标记的训练图像而提高,当只有添加非常有限数量(例如10%)的真实图像数据时,分割性能较低,随着真实数据的数量增多,性能得到提升。当具有专家标注的真实数据数量一定时,添加生成MRI数据的比例越高,分割性能越高。

尽管用少量(例如10%)真实MRI图像训练的基线模型不太准确,但是添加大规模生成MRI数据极大地提高了分割性能。另外,最重要的是在不同比例的混合数据中,本文设计的网络仍然有很高的分割准确率。结果证明了分割算法的性能可以随着更多生成数据的可用而持续提高,这也说明了生成策略的高效性。

表1 使用U-Net计算不同数据扩充策略的分割精度

Table 1 Segmentation accuracy of different data augmentation strategies using U-Net

方法	验证集			测试集		
	DSC	HD95/mm	VR	DSC	HD95/mm	VR
RF+fCRF	0.49±0.17	12.44±10.31	0.58±0.63	0.53±0.20	14.81±10.27	0.59±0.48
U-Net(tMRI)	0.64±0.25	7.01±7.23	0.33±0.30	0.50±0.27	18.59±13.87	0.52±0.30
CycleGAN(t&p MRI)	0.58±0.22	11.21±6.47	0.52±0.52	0.64±0.18	15.37±10.59	0.55±0.98
StyleGAN(t&p MRI)	0.61±0.28	8.52±6.19	0.55±0.58	0.57±0.19	18.03±14.20	0.41±0.34
GAN(t&p MRI)	0.59±0.23	10.87±6.01	0.47±0.37	0.63±0.23	16.11±15.01	0.49±0.36
StarGAN(t&p MRI)	0.66±0.27	7.87±4.37	0.64±0.41	0.60±0.22	20.37±18.07	0.34±0.51
本文方法(t&p MRI)	0.71±0.20	5.72±3.41	0.26±0.17	0.74±0.13	10.31±6.57	0.17±0.18
本文方法(pMRI)	0.62±0.25	7.52±5.81	0.35±0.29	0.72±0.15	12.53±11.38	0.28±0.21

表2 使用RFCN和DFCN计算不同数据扩充策略的分割精度

Table 2 Segmentation accuracy of different data augmentation strategies using RFCN and DFCN

方法	网络	验证集			测试集		
		DSC	HD95/mm	VR	DSC	HD95/mm	VR
U-Net(tMRI)	RFCN	0.63±0.22	9.01±5.23	0.31±0.32	0.51±0.24	23.78±16.90	0.50±0.24
	DFCN	0.67±0.26	8.36±4.73	0.28±0.27	0.58±0.25	24.88±16.23	0.53±0.92
CycleGAN(t&p MRI)	RFCN	0.54±0.32	11.71±5.38	0.42±0.32	0.59±0.20	18.35±12.59	0.56±1.01
	DFCN	0.60±0.28	10.02±5.83	0.33±0.24	0.61±0.18	23.80±17.58	0.38±0.55
StyleGAN(t&p MRI)	RFCN	0.43±0.27	16.52±6.18	0.55±0.52	0.59±0.29	20.03±13.09	0.49±1.26
	DFCN	0.61±0.24	9.85±6.85	0.49±0.37	0.56±0.26	25.30±21.02	0.44±0.38
GAN(t&p MRI)	RFCN	0.50±0.29	11.87±8.00	0.46±0.38	0.55±0.28	22.11±15.87	0.55±0.93
	DFCN	0.52±0.33	12.73±9.09	0.50±0.41	0.54±0.26	23.09±16.01	0.51±0.76
StarGAN(t&p MRI)	RFCN	0.67±0.25	7.56±4.52	0.40±0.41	0.60±0.25	17.37±17.07	0.34±0.47
	DFCN	0.68±0.20	6.46±4.28	0.45±0.44	0.58±0.23	27.87±22.88	0.46±0.51
本文方法(t&p MRI)	RFCN	0.74±0.21	6.83±4.21	0.27±0.21	0.75±0.17	15.31±7.57	0.22±0.20
	DFCN	0.72±0.20	6.53±5.13	0.30±0.27	0.77±0.16	14.28±8.03	0.19±0.17
本文方法(pMRI)	RFCN	0.54±0.24	10.28±5.78	0.37±0.31	0.68±0.20	18.35±12.80	0.29±0.27
	DFCN	0.60±0.21	11.03±5.37	0.45±0.29	0.65±0.18	19.99±11.83	0.23±0.22

表3 不同比例生成数据对分割结果影响(%)

Table 3 Effects of different proportions of generated data on segmentation results (%)

tMRI	FCN(MIoU)				U-Net(MIoU)			
	10% gMRI	25% gMRI	50% gMRI	100% gMRI	10% gMRI	25% gMRI	50% gMRI	100% gMRI
10%	48.75	54.30	59.53	62.82	45.05	53.36	62.53	68.56
25%	50.03	57.96	66.33	70.59	53.20	60.01	67.94	74.87
50%	52.89	64.83	73.55	79.12	56.99	66.38	79.01	80.37
100%	55.27	65.66	73.01	81.42	58.00	68.35	64.05	83.71

FCN:全卷积网络;tMRI:真实MRI图像;gMRI:生成MRI图像;MIoU:平均交并比

最终研究结果显示,CSCGAN实现了CT到MRI的模态转换,并且与目前主流的生成对抗网络相比,包括GAN、StarGAN、CycleGAN、StyleGAN等,本文模型提高了CT到MRI之间的转换效果,且在分割过程中,使用生成的大量伪MRI可以在没有充足专家标注真实数据的情况下提高分割的准确度。

4 结语

本文针对如何利用未配对的CT和MRI数据进行MRI数据的扩充,详细描述了生成对抗网络和混合注意力模型的应用,以CycleGAN模型为主体框架,结合StyleGAN2精细控制生成细节的特点,生成了精确的MRI图像。修改了噪声结构,融入注意力机制减少了肿瘤生成过程中的丢失。本文方法利用了生成对抗网络产生的伪MRI,通过比较所有生成的伪MRI图像的强度直方图和真实MRI图像计算的强度直方图,并且将其放入不同的分割网络中进行判断。最后经过评估,本文模型和其他模型相比,生成的数据在分割网络中,K-L散度最低,有最高的DSC、最低的VR和最小的Hausdorff距离,提升了肿瘤分割网络的准确率,表明在分割网络的应用中,比其他生成方法生成的数据有更高的准确性。同时,将不同比例的生成数据混合到真实数据中,结果表明,生成的数据对分割性能具有增益作用。

【参考文献】

- [1] Zhu J, Park T, Isola P, et al. Unpaired image-to image translation using cycle-consistent adversarial networks[C]//International Conference on Computer Vision, 2017.
- [2] Karras T, Laine S. Analyzing and improving the image quality of StyleGAN[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2019: 8107-8116.
- [3] Li W, Li YF, Qin WJ, et al. Magnetic resonance image (MRI) synthesis from brain computed tomography (CT) images based on deep learning methods for magnetic resonance (MR)-guided radiotherapy[J]. Quant Imaging Med Surg, 2020, 10(6): 1223-1236.
- [4] Yang H, Sun J, Carass A, et al. Unsupervised MR-to-CT synthesis using structure-constrained CycleGAN[J]. IEEE Trans Med Imaging, 2020, 39(12): 4249-4261.
- [5] Al-Masni MA, Al-Antari MA, Park HM, et al. A deep learning model Integrating FRCN and residual convolutional networks for skin lesion segmentation and classification[C]//Proceedings of the 2019 IEEE Eurasia Conference on Biomedical Engineering, Healthcare and Sustain ability. IEEE, 2019.
- [6] Emami H, Dong M, Nejad-Davarani SP, et al. Generating synthetic CTs from magnetic resonance images using generative adversarial networks [J]. Med Phys, 2018, 45(8): 3627-3636.
- [7] Emami H, Dong M, Nejad-Davarani SP, et al. SA-GAN: structure-aware GAN for organ-preserving synthetic CT generation [C]//Medical Image Computing and Computer-Assisted-MICCAI. Intervention Berlin: Springer, 2021: 471-481.
- [8] Qi MK, Li YB, Wu AQ, et al. Multi-sequence MR image-based synthetic CT generation using a generative adversarial network for head and neck MRI-only radiotherapy[J]. Med Phys, 2020, 47(4): 1880-1894.
- [9] Lei Y, Harms J, Wang T, et al. MRI-only based synthetic CT generation using dense cycle consistent generative adversarial networks[J]. Med Phys, 2019, 46(8): 3565-3581.
- [10] Yang H, Sun J, Carass A, et al. Unsupervised MR-to-CT synthesis using structure-constrained CycleGAN[J]. IEEE Trans Med Imaging, 2020, 39(12): 4249-4261.
- [11] Ge YH, Wei DM, Xue Z, et al. Unpaired MR to CT synthesis with explicit structural constrained adversarial learning [C]//IEEE International Symposium on Biomedical Imaging. IEEE, 2019: 1096-1099.
- [12] Zachiu C, De Senneville BD, Moonen CT, et al. Anatomically plausible models and quality assurance criteria for online mono and multi-modal medical image registration[J]. Phys Med Biol, 2018, 63(15): 155016.
- [13] Bano S, Asad M, Fetit AE, et al. XmoNet: a fully convolutional network for cross-modality MR image inference [C]//First International Workshop, PRIME 2018, 2018: 129-137.
- [14] Xiang L, Wang Q, Nie D, et al. Deep embedding convolutional neural network for synthesizing CT image from T1-Weighted MR image[J]. Med Image Anal, 2018, 47: 31-44.
- [15] Wang C, Macnaught G, Papanastasiou G, et al. Unsupervised learning for cross-domain medical imagesynthesis using deformation invariant cycle consistency networks[C]//International Workshop on Simulation and Synthesis in Medical Imaging, 2018: 52-60.
- [16] Hiasa Y. Cross-modality image synthesis from unpaired datausing cyclegan[C]//International Workshop on Simulation and Synthesis in Medical Imaging, 2018: 31-41.
- [17] Wu L, Wang Y, Shao L. Cycle-consistent deep generative hashing for cross-modal retrieval[J]. IEEE Trans Image Process, 2019, 28(4): 1602-1612.
- [18] Xiang L, Li Y, Lin W, et al. Unpaired deep cross-modality synthesis with fast training [C]//Medical Image Computing and Computer Assisted Intervention-MICCAI 2018, 2018: 215-223.
- [19] Zhang R, Pfister T, Li J. Harmonic unpaired image-to-image translation [J]. arXiv preprint, arXiv:1903.06399v1, 2019.
- [20] Huo YK, Xu ZB, Bao SX, et al. Landman, Adversarial synthesis learning enables segmentation without targetmodality ground truth [C]//IEEE International Symposium on Biomedical Imaging, 2018: 1217-1220.
- [21] Ge YH, Xue Z, Cao TY, et al. Unpaired whole-body MR to CT synthesis with correlation coefficient constrained adversarial learning [M]. Image Processing, 2019: 28-35.
- [22] Nie D, Trullo R, Lian J, et al. Medical image synthesis with deep convolutional adversarial networks[J]. IEEE Trans Bio-Med Eng, 2018, 65(12): 2720-2730.
- [23] Cao X, Yang J, Gao Y, et al. Region-adaptive deformable registration of CT/MRI pelvic images via learning-based image synthesis[J]. IEEE Trans Image Process, 2018, 27: 3500-2512.
- [24] Wolterink JM, Dinkla AM, Savenije MH, et al. Deep MR to CT synthesis using unpaired data[J]. arXiv preprint, arXiv:1708.01155, 2017.
- [25] Jiang J, Hu YC, Tyagi N, et al. Tumor-aware, adversarial domain adaptation from CT to MRI for lung cancer segmentation. Medical image computing and computer-assisted intervention [C]//Medical Image Computing and Computer Assisted Intervention Society. Berlin: Springer, 2018.
- [26] Dar S, Yurt M, Karacan L, et al. Image synthesis in multi-contrast MRI with conditional generative adversarial networks[J]. IEEE Trans Med Imaging, 2019, 38(10): 2375-2388.
- [27] Chatsias A, Joyce T, Dharmakumar R, et al. Adversarial image synthesis for unpaired multi-modal cardiac data [C]//International Workshop on Simulation and Synthesis in Medical Imaging. Springer, 2017: 3-13.
- [28] Jin CB, Jung W, Joo S, et al. Deep CT to MR synthesis using paired and unpaired data[J]. Sensors (Basel) 2019; 19(10): 2361.
- [29] Abu-Srhan A, Almallahi I, Abushariah MA, et al. Paired-unpaired unsupervised attention guided GAN with transfer learning for bidirectional brain MR-CT synthesis[J]. Comput Biol Med, 2021, 136 (2): 104763.
- [30] Karras T, Laine S, Aila T. A style-based generator architecture for generative adversarial networks[C]//IEEE Conference on Computer Vision and Pattern Recognition, 2019.
- [31] Cohen JP, Luck M, Honari S. Distribution matching losses can hallucinate features in medical image translation [C]//International Conferenceon Medical Image Computing and Computer-Assisted Intervention. Springer, 2018: 529-537.
- [32] Jiang J, Riyahi Alam S, Chen I, et al. Deep cross-modality (MR-CT) educed distillation learning for cone beam CT lung tumor segmentation [J]. Med Phys, 2021, 48(7): 3702-3713.

(编辑:薛泽玲)