

融合感知损失的深度学习在常规MR图像转换的研究

张泽茹^{1,2}, 李兆同^{1,2}, 刘良友^{1,2}, 高嵩¹, 吴奉梁³

1. 北京大学医学部医学技术研究院, 北京 100191; 2. 北京大学医学人文学院, 北京 100191; 3. 北京大学第三医院骨科, 北京 100191

【摘要】目的:研究在完全无监督的条件下深度神经网络实现常规磁共振图像间相互转换的可行性。**方法:**在循环生成式对抗网络(CycleGAN)中引入感知损失,使网络利用对抗损失学习图像结构信息的同时,结合循环一致性损失和感知损失生成高质量的磁共振图像,并将生成图像与CycleGAN模型以及有监督的CycleGAN模型(S_CycleGAN)生成的图像进行定量比较。**结果:**引入感知损失后的网络生成的图像定量评估值均高于CycleGAN模型生成的图像,生成的 T_1 加权图像(T_1WI)的定量评估值也均高于S_CycleGAN模型生成的 T_1WI ,生成的 T_2 加权图像(T_2WI)与S_CycleGAN模型生成的 T_2WI 的定量评估值相似。**结论:**在CycleGAN中引入感知损失,可以在完全无监督的条件下生成高质量的磁共振图像,进而实现高质量的常规磁共振图像的相互转换。

【关键词】磁共振成像;多模态;图像转换;生成式对抗网络

【中图分类号】R445.2;R318

【文献标志码】A

【文章编号】1005-202X(2021)02-0178-08

Application of deep learning with perceptual loss in conventional MR image translation

ZHANG Zeru^{1,2}, LI Zhaotong^{1,2}, LIU Liangyou^{1,2}, GAO Song¹, WU Fengliang³

1. Institute of Medical Technology, Peking University Health Science Center, Beijing 100191, China; 2. School of Health Humanities, Peking University, Beijing 100191, China; 3. Department of Orthopedics, Peking University Third Hospital, Beijing 100191, China

Abstract: Objective To research the feasibility of using deep neural networks to achieve image-to-image translation on conventional magnetic resonance (MR) images in a completely unsupervised way. **Methods** Perception loss was introduced into cycle generative adversarial network (CycleGAN), so that the proposed network could use the adversarial loss to learn image structure information, and combine cycle consistency loss with perceptual loss to generate high-quality MR image. The generated image was compared quantitatively with those generated by CycleGAN model and supervised CycleGAN model (S_CycleGAN). **Results** The quantitative evaluation showed that the proposed network with the introduction of perceptual loss was superior to CycleGAN model on imaging, and that the evaluation result of the T_1 -weighted image generated by the proposed network was also better than that of the image generated by S_CycleGAN model. However, the evaluation results of the T_2 -weighted images generated by the proposed network and S_CycleGAN model were similar. **Conclusion** The introduction of perceptual loss to CycleGAN can generate high-quality MR images in a completely unsupervised way, and then realize image-to-image translation on high-quality conventional MR images.

Keywords: magnetic resonance imaging; multi-modalities; image translation; generative adversarial network

前言

磁共振成像具有高分辨率软组织图像、病变显

示敏感性高的优点,被广泛应用于临床诊断。不同的脉冲序列在成像时可以获取到不同模态的磁共振(Magnetic Resonance, MR)图像,这些模态图像的对比机制不同,提供的病理和生理信息也不同,因此同时使用多模态的MR图像可以获得更多的病理信息。例如, T_1 (自旋-晶格弛豫)加权图像(T_1 Weighted Image, T_1WI)适用于观察解剖结构,而 T_2 (自旋-自旋弛豫)加权图像(T_2 Weighted Image, T_2WI)突出显示组织中的液体,适用于观察组织病变。同时获取高质量的多模态MR图像比较困难。首先获取需要较长的扫描时间,其次受患者不自主运动影响获取到

【收稿日期】2021-01-04

【基金项目】国家自然科学基金(12075011,82071280);北京市自然科学基金(7202093);西藏自治区重点研发计划(XZ202001ZY0005G)

【作者简介】张泽茹,硕士研究生,研究方向:磁共振成像, E-mail: 2011210185@stu.pku.edu.cn

【通信作者】吴奉梁,博士,副教授,副主任医师,研究方向:脊柱外科, E-mail: wfl79gxy81@163.com

的MR图像可能会出现伪影而不能用于诊断^[1]。因此,通过利用扫描时间相对较短的图像来生成另一模态图像,或通过高质量的某一模态图像生成损坏的另一模态图像对获取多模态MR图像具有潜在的临床应用价值^[2]。

图像到图像转换网络在自然图像上的成功,为多模态医学图像转换提供了很好的参考。生成式对抗网络(Generative Adversarial Network, GAN)通过生成模型和判别模型之间的对抗来实现总目标函数的自动学习,使生成模型输出图像尽可能接近真实图像^[3]。目前已有许多研究将GAN用于多模态医学图像转换,如端到端的三维卷积神经网络^[4]、三维条件生成式对抗网络^[5-7]、全卷积网络的GAN模型^[8]、有监督的GAN^[9]和无监督的GAN^[10-12]等。有监督的模型在不同模态MR图像间的转换表现优于无监督的模型表现^[2],但是成对的MR图像较难获取,因此需要一个通过对未配对MR图像进行学习生成高质量MR图像的模型。鉴于在多数图像处理任务中很难获取到成对的训练数据,Zhu等^[10]利用转换具有循环一致性的特性,提出了无监督的循环生成式对抗网络(CycleGAN),在GAN网络中引入循环一致性损失,将该损失与对抗损失组合进行学习,实现了未配对图像到图像的转换。

本文对CycleGAN模型进行改进,在使用对抗损失保证图像结构信息和循环损失保证循环一致性的同时,引入感知损失提高生成图像的准确性。Welander等^[12]提及到在训练中引入真实图像可以使模型学习到更多的图像信息,且本文使用成对的数据集进行训练,因此本文还搭建了一个训练可监督的CycleGAN模型(S_CycleGAN)与改进后的模型(M_CycleGAN)进行对比。通过对M_CycleGAN模型、S_CycleGAN模型及CycleGAN模型进行定量与可视化分析,研究M_CycleGAN模型无监督、高质量地实现常规MR图像间相互转换的可行性。

1 方法

GAN是由两个子网络 G 和 D 组成的神经网络结构,其中 G 是生成器, D 是鉴别器。 G 旨在从随机向量噪声 z 到目标域中图像 y 的映射,而 D 学习区分生成的图像 $G(z)$ 和真实图像 y 。不同于GAN网络,CycleGAN同时学习两个映射。定义两个图像域 X 和 Y ,如图1a所示,CycleGAN学习从给定图像 $x(x \in X)$ 到输出图像 $y(y \in Y)$ 的映射 $G: x \rightarrow y$ 以及从给定图像 y 到输出图像 x 的映射 $F: y \rightarrow x$ 。在CycleGAN中还包含了两个鉴别器 D_x 和 D_y ,生成器

G 和 F 尽可能生成真实的图像以欺骗鉴别器,而作为一个二分类模型的鉴别器模型, D_x 用于区分真实图像 x 和预测图像 $F(y)$, D_y 用于区分真实图像 y 和预测图像 $G(x)$ 。本文提出的M_CycleGAN在CycleGAN的基础上引入感知损失,通过一个已训练的网络分别提取重建图像和真实图像的高层特征信息,并通过计算两幅图像提取特征的差异学习两幅图像之间的高级语义信息。S_CycleGAN在CycleGAN的基础上引入一个L1损失,通过计算生成图像与真实图像之间的L1损失实现模型的可监督训练。

本文中模型的损失函数包含4种:用于将生成图像的数据分布与目标图像的数据分布相匹配的对抗损失;用于防止学习到的映射函数 G 和 F 之间矛盾的循环一致性损失;用于提高图像逼真度的重建图像与真实图像之间的感知损失;实现模型可监督的生成图像和真实图像之间的L1损失。

1.1 对抗损失

将对抗损失用于映射函数 $G: X \rightarrow Y$,则对映射函数 G 和它的鉴别器 D_y 而言,其损失函数表示为:

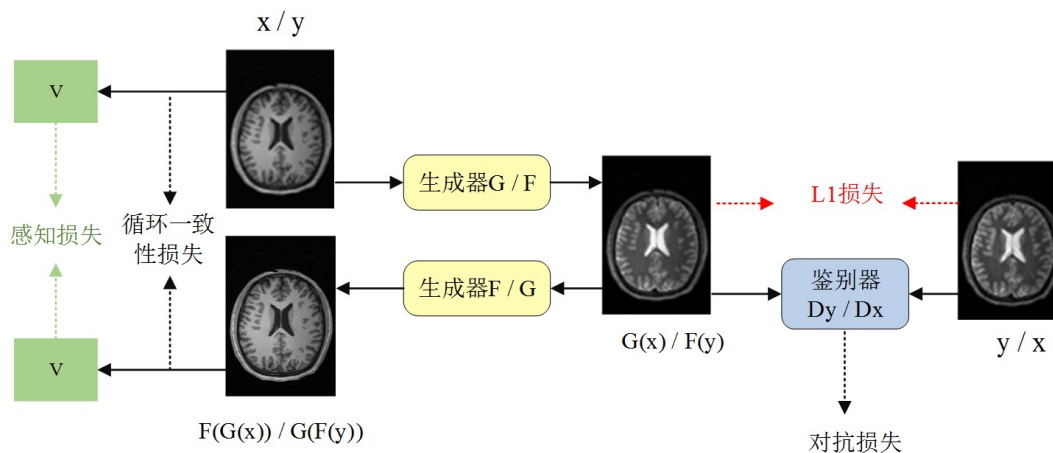
$$L_{\text{GAN}}(G, D_y, X, Y) = E_{y' \sim p_{\text{data}}(y)} [\log D_y(y')] + E_{x \sim p_{\text{data}}(x)} [1 - \log D_y(G(x))] \quad (1)$$

其中,生成器 G 用来生成与 Y 中的真实图像尽可能一致的图像以迷惑 D_y ,而 D_y 用来最大努力地区分生成图像 $G(x)$ 和真实图像 y 。类似博弈过程, G 旨在将该损失最小化, D_y 则旨在将该损失最大化,训练目标为 $\min_G \max_{D_y} L_{\text{GAN}}(G, D_y, X, Y)$ 。将对抗损失也应用于映射函数 $F: Y \rightarrow X$,则其和对应鉴别器 D_x 的损失函数为 $L_{\text{GAN}}(F, D_x, X, Y)$ 。

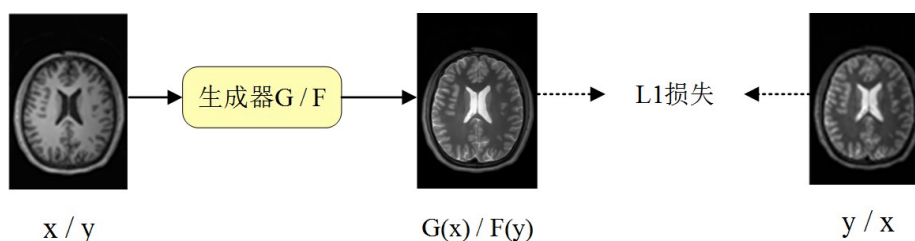
1.2 循环一致性损失

理论上,通过对抗训练可以学习到映射函数 G 和映射函数 F ,两个映射分别通过源域 X 和 Y 的输入产生与目标域 Y 和 X 相匹配的分布。但由于 G 和 F 是随机的函数,有无限多个映射函数 G 和 F 可以使其输出数据的分布与目标域的分布相同,因此只使用对抗训练并不能保证单个输入 x 和输出 y 的配对有意义,且还有可能导致模型崩溃^[10]。

在执行转换任务的过程中,可以观察到转换满足循环一致的特性,即域 X 中的图像 x 通过映射函数 G 生成 $G(x)$,通过转换循环可以再返回到原来的图像,即 $x \rightarrow G(x) \rightarrow F(G(x)) \approx x$ 。类似地,域 Y 中的图像 y 通过映射函数 F 生成 $F(y)$,通过循环一致也可以返回到原来的图像,即 $y \rightarrow F(y) \rightarrow G(F(y)) \approx y$ 。因此为了减少映射函



a: CycleGAN 模型结构示意图(黑色为 CycleGAN 模型结构示意图, 绿色为 M_CycleGAN 模型的改进部分, 红色为 S_CycleGAN 模型的改进部分)



b: Generator 模型结构示意图

图1 模型结构示意图

Fig.1 Network architectures

数 G 和 F 的选择空间, 引入循环一致性损失。这一损失的定义为:

$$L_{\text{cyc}}(G, F) = E_{y \sim p_{\text{data}}(y)} [\|G(F(y)) - y\|_1] + E_{x \sim p_{\text{data}}(x)} [\|F(G(x)) - x\|_1] \quad (2)$$

1.3 感知损失

上述提及到的循环一致性损失使用的逐像素损失在计算仅差一个像素的两幅相同图像时, 尽管两幅图像的感知相似, 也会出现很大差异。这是由于逐像素损失不能捕捉生成图像和真实图像之间的感知差异。研究表明在网络训练中结合感知损失可以在计算机视觉任务中产生更真实的视觉效果^[13-16]。不同于像素级损失函数依赖于像素间的差异, 感知损失依赖于高级语义特征的差异, 一般这些高级特征通过预训练的网络进行提取。本文使用在 ImageNet 上预训练的 VGG16^[17] 提取特征图。将感知损失用于重建图像和真实图像, 结合循环一致性损失进一步缩小映射函数 G 和 F 的选择空间, 则其损失函数为:

$$L_{\text{perc}}(G, F) = E_{x, y \sim p_{\text{data}}(x, y)} [\|V(F(G(x))) - V(x)\|_1 + \|V(G(F(y))) - V(y)\|_1] \quad (3)$$

其中, V 表示 VGG16 中第二次最大池化前的特征映射图。

1.4 L1 损失

由于生成器 G 和 F 在生成图像迷惑鉴别器的同时, 也要求产生高质量的图像。且在大多数情况下, L1 损失能够捕捉低频信息、保证图像低层特征的准确性。因此本文在 S_CycleGAN 中引入 L1 损失项来实现 CycleGAN 模型的可监督训练。映射函数 G 的 L_1 损失项定义为:

$$L_{L1}(G) = E_{x \sim p_{\text{data}}(x), y \sim p_{\text{data}}(y)} [\|G(x) - y\|_1] \quad (4)$$

映射函数 F 的 L_1 损失项定义为:

$$L_{L1}(F) = E_{x \sim p_{\text{data}}(x), y \sim p_{\text{data}}(y)} [\|F(y) - x\|_1] \quad (5)$$

1.5 总目标函数

CycleGAN 的总损失函数表达式为:

$$L(G, F, D_X, D_Y) = L_{\text{GAN}}(G, D_Y, X, Y) + L_{\text{GAN}}(F, D_X, X, Y) + \lambda L_{\text{cyc}}(G, F) \quad (6)$$

其中, λ 用来控制循环损失的权重。

S_CycleGAN 的总损失函数表达式为:

$$L(G, F, D_X, D_Y) = L_{\text{GAN}}(G, D_Y, X, Y) + L_{\text{GAN}}(F, D_X, X, Y) + \lambda_1 L_{\text{cyc}}(G, F) + \lambda_2 L_{L1}(F) + \lambda_2 L_{L1}(G) \quad (7)$$

其中, λ_1 用来控制循环损失的权重, λ_2 用来控制 L1 损失的权重。

M_CycleGAN 的总损失函数表达式为:

$$L(G, F, D_x, D_y) = L_{\text{GAN}}(G, D_y, X, Y) + L_{\text{GAN}}(F, D_x, X, Y) + \lambda_1 L_{\text{cyc}}(G, F) + \lambda_2 L_{\text{perc}}(G, F) \quad (8)$$

其中, λ_1 用来控制循环损失的权重, λ_2 用来控制感知损失的权重与对抗训练中相互竞争的思想相同, 生成器 G 和 F 旨在将损失最小化, 鉴别器 D_x 和 D_y 旨在将损失最大化。因此上述 3 种网络的总目标函数均为:

$$G^*, F^* = \arg \min_{G, F} \max_{D_x, D_y} L(G, F, D_x, D_y) \quad (9)$$

为了研究对抗性和循环一致性损失对模型的影响, 本文还实现了 Generator 模型。它由 CycleGAN 的生成器组成, 仅使用 L1 损失在可监督的方式进行训练。其损失函数如式(4)和式(5)所示。

2 跨模态 MR 图像生成

2.1 网络结构

本文使用 M_CycleGAN、S_CycleGAN、CycleGAN 以及 Generator 模型生成跨模态 MR 图像。图 1a 中的 S_CycleGAN、M_CycleGAN 与 CycleGAN 模型均由两个相同结构的生成器和两个相同结构的

鉴别器组成。图 1b 中 Generator 模型中的生成器结构与 CycleGAN 模型中的生成器结构相同。

2.1.1 生成器结构 在 T_1 WI 和 T_2 WI 相互转换研究中, 由于输入图像和输出图像大小相同, 因此生成器使用了上采样层数量和下采样层数量相同的编码-解码网络。另一方面先进行下采样再进行上采样的方法可以在计算代价相同的情况下, 使用更大的网络。

生成器的结构中有 4 个卷积层、9 个残差块以及 2 个转置卷积层, 如图 2 所示。本网络中没有使用任何池化层, 而是参照 Radford 等^[18]提出的网络及指导, 使用带有步长的卷积层来代替池化层以实现网络中的图像上采样和下采样。使用残差连接能够让网络更容易地学习确定的函数, 减少数据的丢失和损耗、保护输入信息的完整性^[19], 因此在网络中加入了残差块以在保证信息完整性的同时增加网络深度, 其中残差块的结构为两个使用 1×1 卷积核的卷积层。由于图像转换的生成结果主要依赖于对应的输入图像实例, 因此使用实例归一化对数据进行归一化, 以保持每张图像实例之间的独立。

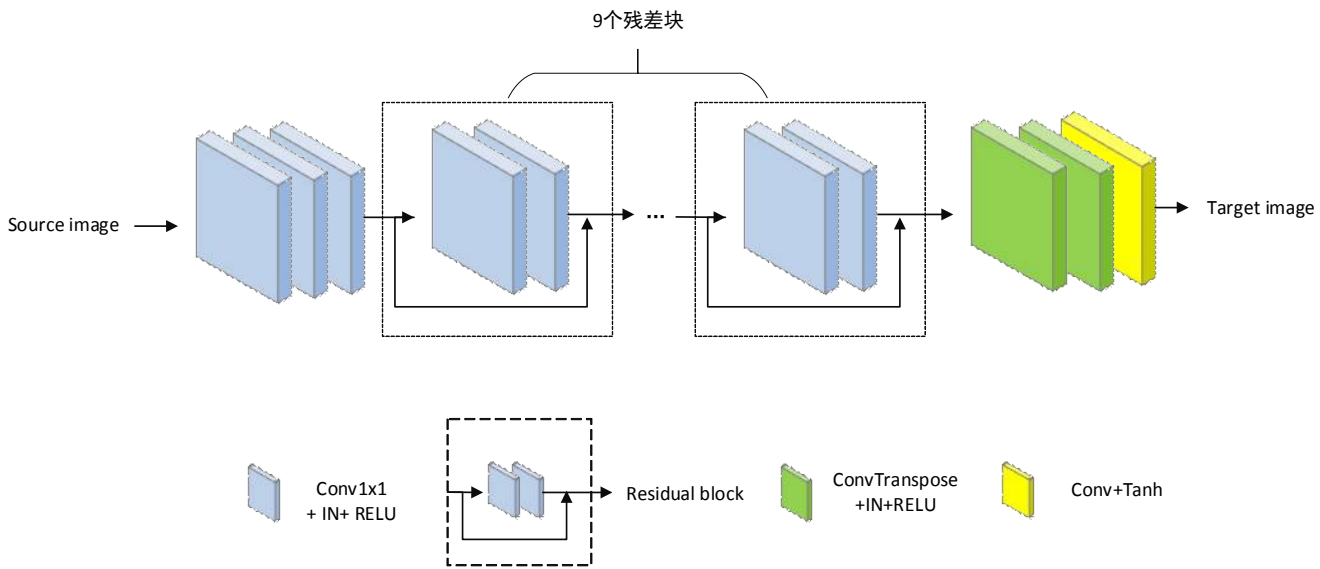


图2 生成器结构示意图

Fig.2 Network architecture of generator

2.1.2 鉴别器结构 在 GANs 模型中, 生成器生成的图像被用作鉴别器的输入之一与真实图像进行定量比较, 以优化生成器的性能。鉴别器的目的是将生成器生成的图像判别为假的即标签为 0, 真实的图像判别为真的即标签为 1, 然后将结果反馈给生成器。根据实验证明在图像生成问题中使用 L1 损失和 L2 损失产生的结果虽然会比较模糊, 但是两者均可以准确地捕捉低频信息^[9,20]。在 CycleGAN 网络中通过

计算重建图像与输入图像的 L1 损失来实现循环一致性, 所以在 CycleGAN 网络中可以使用循环一致性损失来捕捉图像的低频特征, 使用 PatchGAN^[9] 作为鉴别器来获取图像的高频信息。PatchGAN 鉴别器只对块规模的结构进行惩罚, 即判别一个图像中 $N \times N$ 大小的块是真的还是假的, 然后将这一鉴别器对整个图像进行卷积, 将结果平均化后作为鉴别器最后的输出。这一特性将注意力集中到图像的局部块

上,满足先前 GAN 鉴别器对高频信息进行建模的需求,且使用 PatchGAN 比使用整幅图像的鉴别器所需参数更少、训练时间更短。在本文中,使用 70×70 的 PatchGAN 作为鉴别器网络,判别 70×70 重叠的图像

块是真的还是假的。如图 3 所示为鉴别器的结构示意图。不同于生成器,鉴别器中使用斜率为 0.2 的带泄露修正线性单元(LeakyReLU, LReLU)进行非线性激活。

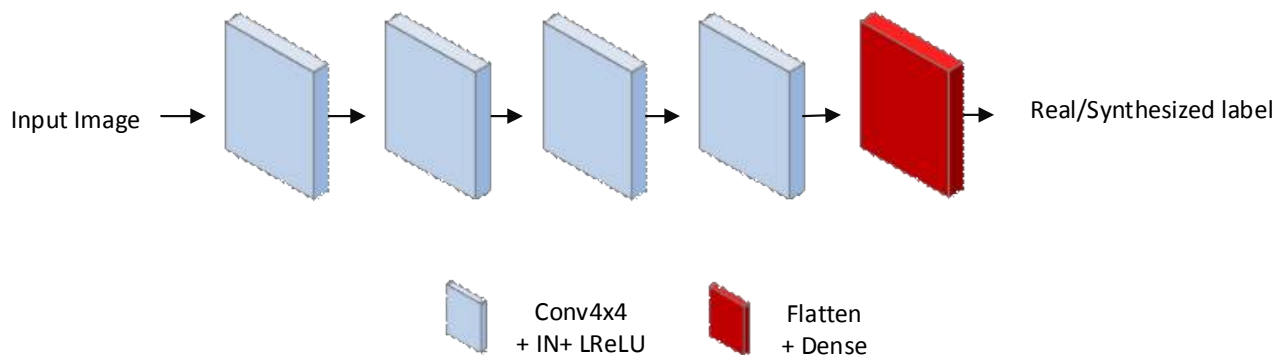


图3 鉴别器结构示意图

Fig.3 Network architecture of discriminator

2.2 实验细节

本文的模型使用 Nvidia GeForce GTX 1060 GPU 进行训练。训练的数据集由 Human Connectome project(<https://db.humanconnectome.org/>)提供。本文使用了 32 位受试者成对的 T_1WI 和 T_2WI (但需要注意的是 CycleGAN 网络可以使用未配对的数据进行训练)。其中,将 26 位受试者的数据用于训练,剩余 6 位受试者的数据用于测试。

数据处理主要分为三步。第一步,使用 ITK-SNAP 软件对每位受试者的成对数据进行配准,使配对的图像大小相同、位置相同。第二步,由于本文的模型只能支持二维的输入图像,所以对配准后的数据分别沿着 x 轴、y 轴、z 轴进行切割,其中水平面选择 30 个切片,矢状面选择 20 个切片,冠状面选择 20 个切片。因此训练集的每个域有 1 820 幅图像,测试集的每个域有 420 幅图像。第三步,图像大小调整。数据集集中的三维数据大小为 256×256×176,切割后的数据大小分别为 256×256、256×176、256×176,因此对大小为 256×176 的图像使用双线性插值调整图像大小,将所有的图像大小统一为 256×256。为了使模型更容易收敛,在进行训练前,需要对数据进行归一化。由于 MR 图像是 8 bit 深度图像,因此使用公式 $X = \frac{\bar{X}}{127.5} - 1$ 将数据统一映射到 -1 到 1 的区间上。

3 个 CycleGAN 模型中所有的权重均以高斯分布 $N(0, 0.02)$ 进行初始化,且设循环一致性损失惩罚系数均为 10.0。设 M_CycleGAN 模型中 λ_2 为 10.0 对真实图像和重建图像的感知损失进行正则化,设 S_CycleGAN 模型中的 λ_2 为 10.0 对生成图像与真实

图像间的 L1 损失进行正则化。3 个模型都是用 Adam 优化器进行训练,其中用以控制权重分配的系数 β_1 设置为 0.5,用以控制之前梯度平方影响的系数 β_2 设置为 0.999。因为本文在模型中使用实例归一化,所以将 batchsize 设置为 1。将生成器的学习速率和鉴别器的学习速率初始设置为 0.000 2,在训练的前 100 个周期中保持学习速率不变,然后在后 100 个周期中学习速率逐渐衰减到 0。在每一周期循环中将生成器迭代次数和鉴别器的迭代次数均设置为 1,使生成器和鉴别器进行交替训练。对于 Generator 模型,仅使用 L1 损失进行训练,其余参数设置与上述 CycleGAN 模型参数一致。

2.3 评估准则

使用归一化均方根误差(Normalized Root Mean-Squared Error, NRMSE)、峰值信噪比(Peak Signal-to-Noise Ratio, PSNR)、结构相似性(Structural Similarity Index, SSIM)^[21]、互信息(Mutual Information, MI)^[22]分别对模型测试的结果进行定量评估。PSNR、SSIM 以及 MI 的值越高,NRMSE 值越小表示生成的图像失真越小。当生成图像与真实图像之间的误差很小的情况下,很难对生成图像继续进行视觉评估,因此本文计算两幅图像之间每个像素的绝对差值并将其可视化以实现视觉评估。

3 结果与讨论

表 1 列出了 M_CycleGAN、S_CycleGAN、CycleGAN 以及 Generator 在测试集上的定量评估结果。Generator 模型在所有定量测量中表现最佳。与定量分析结果不同,图 4 的误差可视化图表示

M_CycleGAN 模型生成的 T_1 WI 和 T_2 WI 的视觉效果表现最佳。这是因为 Generator 模型仅使用 L1 损失进行监督训练使模型倾向于生成表面光滑、不真实的 MR 图像,而其余 3 个 GAN 模型利用对抗损失使生成的 MR 图像更加真实。误差可视图和定量评估结果均显示所有模型在生成 T_2 WI 的任务中表现优于生成 T_1 WI 的任务。这是因为在生成 T_1 WI 的时候模型倾向于学习结构信息,而忽视了图像亮度值,使生成的 T_1 WI 平均亮度高于真实 T_1 WI 的平均亮度。

S_CycleGAN 的所有定量测量结果均优于 CycleGAN 模型,但在可视化误差图中,S_CycleGAN 模型生成的 T_1 WI 误差要大于 CycleGAN 模型生成的

表 1 模型定量评估结果

Tab.1 Quantitative evaluation results of the models

模型		PSNR/dB	SSIM	NRMSE	MI
M_CycleGAN	T1->T2	23.983 6	0.682 9	0.305 3	4.967 6
	T2->T1	23.521 5	0.756 7	0.290 4	5.132 8
S_CycleGAN	T1->T2	24.048 4	0.687 5	0.298 0	4.960 5
	T2->T1	23.145 1	0.740 2	0.301 2	5.158 8
CycleGAN	T1->T2	23.173 1	0.571 1	0.327 4	4.947 0
	T2->T1	22.770 9	0.727 3	0.314 7	5.115 6
Generator	T1->T2	25.017 3	0.739 7	0.266 6	4.970 8
	T2->T1	24.064 6	0.778 8	0.269 3	5.170 6

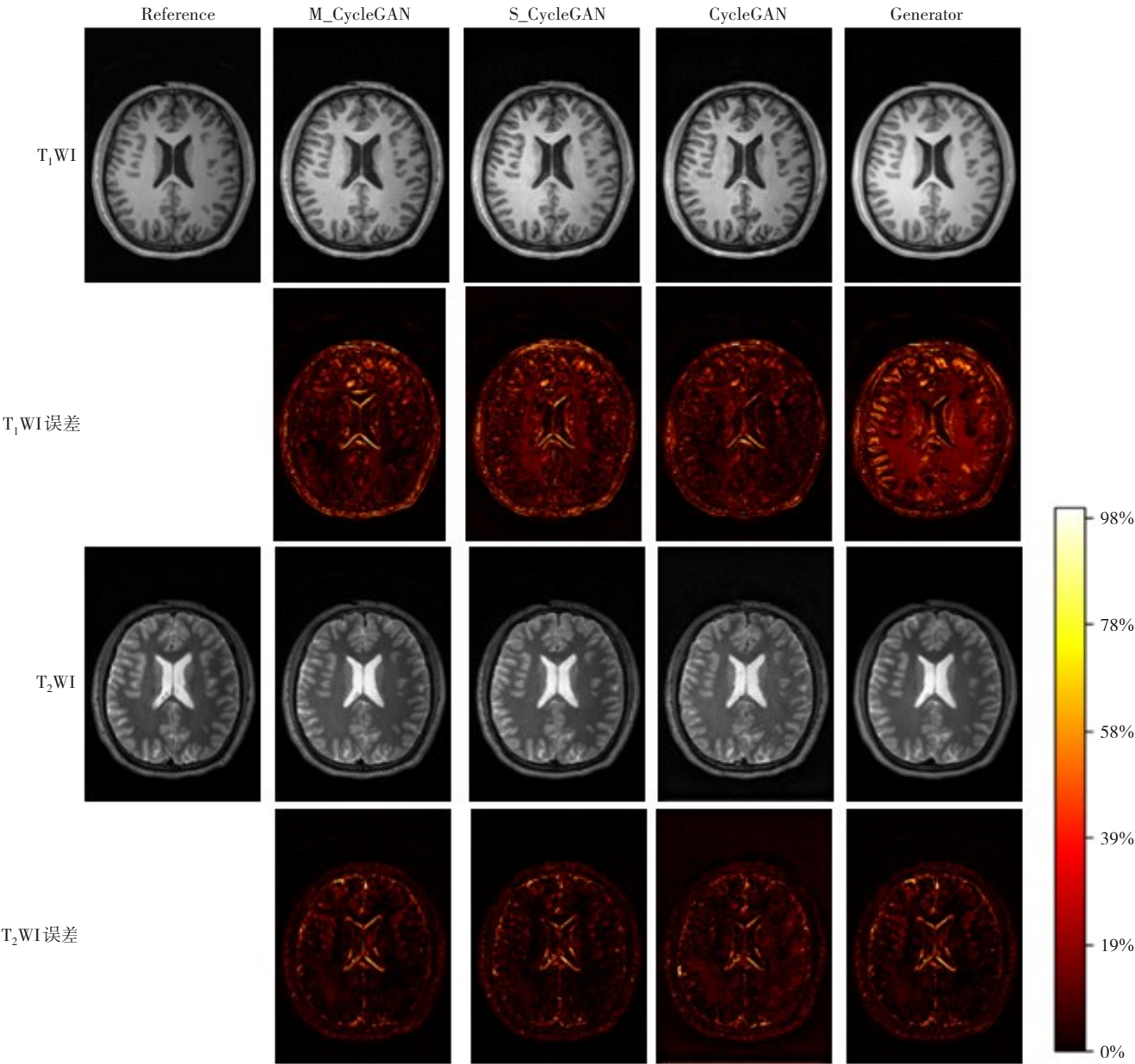


图 4 评估模型的生成图像

Fig.4 Evaluation of images generated by different models

T_1 WI 误差。这是因为与 Generator 模型一样, S_CycleGAN 模型使用了 L1 损失实现 CycleGAN 模型的可监督训练,而 L1 损失是逐像素的计算方法,使生成的 MR 图像更加光滑。

M_CycleGAN 模型的定量评估和可视化评估的结果均优于 CycleGAN 模型。这是因为 CycleGAN 模型仅使用 L1 损失学习重建图像和真实图像之间的误差,而 L1 损失是像素级别的损失计算方式,忽略了图像的高层特征信息, M_CycleGAN 模型在此基础上引入感知损失,通过一个已训练的 VGG16 网络分别提取重建图像和真实图像的高层特征信息,以结构级别的损失计算方式学习两幅图像之间的高级语义信息,减少了生成器映射函数的选择空间,生成与真实图像尽可能一致的 MR 图像。

Dar 等^[2]表明有监督的模型由于包含更多的参考图像信息,性能要优于无监督模型。本文中实现了一个有监督的 CycleGAN 模型(S_CycleGAN)用来与 M_CycleGAN 模型进行比较。在生成 T_1 WI 的任务中,除了 MI 值外, M_CycleGAN 模型的定量评估结果均优于 S_CycleGAN 模型(PSNR 高 0.376 4 dB, SSIM 高 0.016 5, NRMSE 低 0.010 8, MI 低 0.026 0)。在生成 T_2 WI 的任务中, M_CycleGAN 模型与 S_CycleGAN 模型定量评估结果相近(PSNR 相差 0.064 8 dB, SSIM 相差 0.004 6, NRMSE 相差 0.007 3, MI 相差 0.007 1)。在可视化结果中 M_CycleGAN 模型生成的 T_1 WI 和 T_2 WI 视觉效果均优于 S_CycleGAN 模型。这表明 M_CycleGAN 模型可以在完全无监督的条件下达到与 S_CycleGAN 模型相近甚至更优的效果。

Yang 等^[5]、Yu 等^[6]以及 Nie 等^[8]在研究跨模态医学图像转换时,均使用了配对的数据进行监督训练,但获取配对的多模态医学图像比较困难。Zhang 等^[4]提出了实现未配对 CT 图像和 MR 图像转换的网络,但是该网络为了减少生成图像的几何畸变,使用了带有语义标签的图像作为训练目标,不能实现完全无监督转换。Welander 等^[12]为了对基于无监督 CycleGAN 模型和 UNIT 模型的多模态 MR 图像转换进行评估,使用了配对的 MR 图像进行训练。本文提出的方法是在完全无监督的条件下实现对未配对常规 MR 图像的转换,为了将训练结果和真实结果进行比较,本文使用无监督的方式对配对的常规 MR 图像进行了训练,但使用配对数据对训练结果是否存在影响还需要进一步的探索。

Dar 等^[2]在多模态 MR 图像转换的研究中提及到一个受试者连续的磁共振图像之间存在一定的联

系,本文只使用了二维的深度神经网络对图像进行学习,丢失了连续图像之间的空间信息,因此可以对三维的深度神经网络在三维常规 MR 图像相互转换的可行性进行进一步研究。

4 总结

本文对 CycleGAN 模型进行改进,引入感知损失,使模型在学习重建图像和真实图像之间的底层结构信息的同时,学习高级特征信息。模型利用循环一致性损失和感知损失实现未配对数据训练中的一对一映射,并使生成的 MR 图像与真实图像更加相似。该方法可以实现未配对多模态 MR 图像的转换进而获取多模态 MR 图像信息,并通过生成新的 MR 图像数据来进行数据增强进一步实现图像分割和图像配准等任务。

【参考文献】

- [1] KRUPA K, BEKIESIŃSKA-FIGATOWSKA M. Artifacts in magnetic resonance imaging[J]. Pol J Radiol, 2015, 80(1): 93-106.
- [2] DAR S U, 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.
- [3] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks[J]. NIPS, 2014, 3: 2672-2680.
- [4] ZHANG Z, YANG L, ZHENG Y. Translating and segmenting multimodal medical volumes with cycle-and shape-consistency generative adversarial network [C]//The IEEE Conference on Computer Vision and Pattern Recognition. Munich: IEEE, 2018: 9242-9251.
- [5] YANG Q Y, LI N N, ZHAO Z X, et al. MRI cross-modality image-to-image translation[J]. Sci Rep, 2020, 10(1): 1-18.
- [6] YU B, ZHOU L, WANG L, et al. 3D cGAN based cross-modality MR image synthesis for brain tumor segmentation [C]//2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). Washington: IEEE, 2018: 626-630.
- [7] MIRZA M, OSINDERO S. Conditional generative adversarial nets[J]. arXiv preprint arXiv: 1411.1784, 2014.
- [8] NIE D, TRULLO R, LIAN J, et al. Medical image synthesis with deep convolutional adversarial networks[J]. IEEE Trans Biomed Eng, 2018, 65(12): 2720-2730.
- [9] ISOLA P, ZHU J Y, ZHOU T, et al. Image-to-image translation with conditional adversarial networks [C]//30th IEEE Conference on Computer Vision and Pattern Recognition. Hawaii: CVPR, 2017: 5967-5976.
- [10] ZHU J Y, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks [C]//International Conference on Computer Vision. Italy: IEEE, 2017: 2242-2251.
- [11] LIU M Y, BREUEL T, KAUTZ J. Unsupervised image-to-image translation networks[C]//Advances in Neural Information Processing Systems. Long Beach: CA, 2017: 701-709.
- [12] WELANDER P, KARLSSON S, EKLUND A. Generative adversarial networks for image-to-image translation on multi-contrast MR images -a comparison of CycleGAN and UNIT [J]. arXiv preprint arXiv:

1806.07777, 2018.

[13] GATYS L A, ECKER A S, BETHGE M. Texture synthesis using convolutional neural networks[J]. NIPS, 2015, 28: 262-270.

[14] GATYS L, ECKER A, BETHGE M. A neural algorithm of artistic style [J]. J Vision, 2016, 16(12): 326.

[15] GATYS L A, ECKER A S, BETHGE M. Image style transfer using convolutional neural networks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 2414-2423.

[16] JOHNSON J, ALAHI A, LI F F. Perceptual losses for real-time style transfer and super-resolution [C]//Computer Vision-ECCV. Cham: Springer, 2016: 694-711.

[17] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J]. arXiv preprint arXiv: 1409.1556, 2014.

[18] RADFORD A, METZ L, CHINTALA S. Unsupervised representation learning with deep convolutional generative adversarial networks[J]. arXiv preprint arXiv: 1511.06434, 2015.

[19] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016: 770-778.

[20] LARSEN A B, SONDERBY S K, LAROCHELLE H, et al. Autoencoding beyond pixels using a learned similarity metric [C]//International Conference on Machine Learning. New York: PMLR, 2016: 1558-1566.

[21] WANG Z, 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] PLUIM J P, MAINTZ J, VIERGEVER M A. Mutual-information-based registration of medical images: a survey [J]. IEEE Trans Med Imaging, 2003, 22(8): 986-1004.

(编辑:陈丽霞)