

医学影像跨模态生成方法综述

花芸^{1,2}, 康敏诗¹, 刘盼^{1,2}, 郭华源^{1,2}, 金晓宇³, 李轶玮¹, 何昆仑^{1,2}

¹解放军总医院医学创新研究部, 北京 100853; ²解放军总医院医疗大数据应用技术国家工程研究中心, 医学工程军事科研重点实验室, 北京 100853; ³北京航空航天大学, 北京 100083

摘要:医学人工智能研究快速发展, 基于深度学习的图像生成算法已成为研究的热点之一。本文旨在综述自回归模型、变分自编码器、生成式对抗网络和扩散模型这4种主流图像生成算法的发展现状, 并通过计算机断层扫描、磁共振成像和计算机断层血管造影这3种模态分析生成式模型在医学多模态影像转换中的应用。生成模型在医学影像领域不仅具有广阔的应用前景, 还有巨大的价值潜力。

关键词:深度学习; 图像模态; 医学影像; 人工智能; 神经网络

中图分类号:R445; TP391.7

文献标志码:A

文章编号:2095-5227(2025)02-0153-08

DOI: 10.12435/j.issn.2095-5227.24070104

引用本文:花芸, 康敏诗, 刘盼, 等. 医学影像跨模态生成方法综述 [J]. 解放军医学院学报, 2025, 46 (2): 153-160.

Review of medical imaging cross-modal generation methods

HUA Yun^{1,2}, KANG Minshi¹, LIU Pan^{1,2}, GUO Huayuan^{1,2}, JIN Xiaoyu³, LI Yiwei¹, HE Kunlun^{1,2}

¹Medical Innovation Research Department of PLA General Hospital, Beijing 100853, China; ²National Engineering Research Center for Medical Big Data Application Technology, Medical Engineering Military Research Key Laboratory of PLA General Hospital, Beijing 100853, China; ³Beihang University, Beijing 100083, China

Corresponding author: HE Kunlun. Email: kunlunhe@plagh.org

Abstract: In the rapidly developing field of medical artificial intelligence, image generation algorithm based on deep learning has become one of the research hotspots. This paper aims to review the status quos of four major image generation algorithms, namely autoregressive model, variational autoencoder, generative adversarial network and diffusion model, and analyze the application of generative model in medical multimodal image conversion from three modes: computed tomography, magnetic resonance imaging and computed tomography angiography. The generative model not only has broad application prospects in the field of medical imaging, but also has great value potential.

Keywords: deep learning; imaging modalities; medical imaging; artificial intelligence; neural networks

Cited as: Hua Y, Kang MSH, Liu P, et al. Review of medical imaging cross-modal generation methods[J]. Acad J Chin PLA Med Sch, 2025, 46(2): 153-160.

生成模型不仅能深入理解数据的分布, 还能基于此创造出全新的数据样本, 已广泛应用于图像生成、音乐创作、文本生成等领域, 在机器学习领域扮演着至关重要的角色。近年来, 生成模型在医学领域展现出巨大潜力, 尤其在医学影像去噪、超分辨率增强、心电图预测和肿瘤识别等方面。生成模型利用无标注数据集进行训练, 提高数据使用效率的同时, 还显著增强了模型的泛化性和迁移性, 为医疗诊断等奠定了坚实基础。

1 基于深度学习的图像生成算法

图像生成是计算机视觉中的基本问题, 也是重要的研究方向。目前该技术的主要应用方向为图像上色、图像风格转换、图像修复、超分辨率图像生成等。基于深度学习的图像生成算法主要有4类: (1)基于自回归模型的图像生成算法 (autoregressive models); (2)基于变分自编码器的图像生成算法 (variational auto-encoder, VAE); (3)基于生成对抗网络的图像生成算法 (generative adversarial networks, GAN); (4)基于扩散模型的图像生成算法 (diffusion models)。图1以时间轴的方式展示了生成模型的发展历程^[1-4], 从cGAN模型到 Cola-diffusion模型, 反映了人工智能模型在图像生成领域的进步。表1概括了4种图像生成算法的基

收稿日期: 2024-07-01

基金项目: 新一代人工智能国家科技重大专项资助(2021ZD0140408)

第一作者: 花芸, 硕士, 工程师。Email: huayun@301hospital.com.cn

通信作者: 何昆仑, 博士, 主任医师, 教授, 博士生导师。Email: kunlunhe@plagh.org

础架构、训练方法、损失函数以及各自的优势和局限。表2归纳了当前流行的图像生成产品分析。

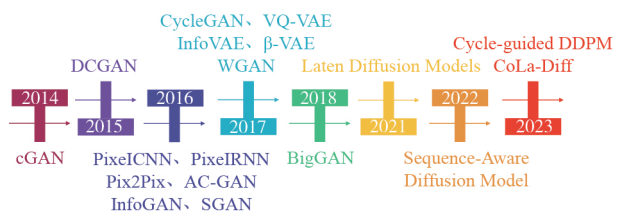


图1 影像生成模型的简要年表

Fig. 1 A brief chronology for image generation models

1.1 基于自回归模型的图像生成算法

基于自回归模型的图像生成算法，主要思想是对顺序的像素点的值进行预测，以前面生成的像素点作为参考，得到后面生成的像素点预测值，然后将所有像素点的概率相乘得到生成图像的分布。主要公式如下：

$$p(x) = \prod_{i=1}^N p(x_i | x_{<i})$$

其中， x 表示整个图像， x_i 表示第*i*个像素点，

表1 基于深度学习图像生成算法总结

Tab. 1 A comprehensive summary of deep learning-based image generation algorithm					
方法	基础架构	训练过程	损失函数	优点	缺点
自回归模型 ^[5]	卷积神经网络	SGD/Adam,增加随机噪声并迭代参数	重构损失+KL散度	训练过程稳定、易评估模型	训练慢、资源需求高、生成质量低
变分自编码器 ^[6]	编码器和解码器架构	优化编码器和解码器参数	重构损失+KL散度	易实现、能并行生成图像	模型评估困难、生成图像质量低
生成对抗网络 ^[7]	生成器和判别器架构	生成器和判别器对抗训练	对抗损失+重构损失	生成图像清晰、能高效并行	训练不稳定、资源需求相对高
扩散模型 ^[8]	扩散过程模型	噪声逐步添加与去除，逆向扩散生成数据	去噪损失	生成图像质量高、多样性强	训练速度慢、资源需求高

表2 图像生成产品对比分析

Tab. 2 Comparative analysis of the image-generated products					
产品	主研单位	FID	基础模型	技术特点	应用功能
Parti	谷歌	7.23	自回归模型,结合 ViT-VQGAN 技术	利用视觉变换器编码图像为离散标记序列,实现高质量图像重建	文本-图像生成;
Imagen ^[9]	谷歌	7.27	Transformer模型和扩散模型结合	结合语言理解和高保真图像生成能力	文本-图像生成;根据文本进行图像编辑
Imagen 2	谷歌	-	扩散模型	生成逼真图像,与用户提示高度契合,支持多语言	文本-图像生成
Make-A- Scee	Meta AI	7.55	-	通过用户草图和文本输入,生成针对性图像	图像生成
DALL-E	OpenAI	17.89	自回归模型	-	文本-图像生成
DALL-E 2 ^[10]	OpenAI	10.39	自回归模型和扩散模型结合	生成高分辨率图像	文本-图像生成
DALL-E 3	OpenAI	-	集成 ChatGPT 的多模态模型	捕捉细微差异,生成连贯高质量图像,增强交互功能	文本-图像生成
DALL-E Mini	CRAFT	-	自回归模型	轻量级版本,适用于资源受限环境	文本-图像生成
GLIDE	OpenAI	12.24	扩散模型	-	文本-图像生成;图像编辑
Midjourney	Midjourney	-	-	相对追求逼真和对文本的精准解析,更加追求艺术感	文本-图像生成
Stable Diffusion	Stability AI	12.63	扩散模型	高效且轻量,可以在消费级 GPU 上运行	文本-图像生成;风格转换
Stable Diffusion XL	Stability AI	-	扩散模型	生成高分辨率图像,支持多样化风格和复杂渲染	文本-图像生成
文心·一格 ^[11]	百度	-	自回归模型	支持中文文本	文本-图像生成
NUWA	微软研究院	-	自回归模型和扩散模型结合	高质量的图像生成	文本-图像生成;图像编辑
NUWA-XL	微软研究院	-	多模态生成模型	能够基于文本和图像信息生成高质量、多样化长视频	文本转图像;视频生成/剪辑

$x_{<i}$ 表示在 x_i 之前的所有像素点， N 是图像中像素点的总数。

1.2 基于变分自编码器的图像生成算法

基于变分自编码器的图像生成算法由Kingma和Welling^[6]在2013年首次提出。由于其出色的生成能力和灵活性，VAE在深度学习领域迅速获得了广泛关注。其主要思想是用编码器和解码器以及变分推断的方法学习隐空间到真实图像空间的映射，从而完成图像的生成。其损失函数如下：

$$L = -E_{q(z|x)}[\log p(x|z)] + KL(q(z|x)||p(z))$$

其中第1项是重参数化后的期望重构损失，第2项是隐变量分布与先验分布之间的Kullback-Leibler (KL)散度。

1.3 基于生成对抗网络的图像生成算法

生成式对抗网络是一种强大的深度学习模型，由Goodfellow等^[7]在2014年首次提出。GAN通过引入两个相互竞争的生成器和判别器来生成数据。GAN的核心思想是模拟一个生成过程，生成器接收随机噪声向量 z 作为输入，这个向量通常为均匀分布或高斯分布。生成器网络通过学习将低维的噪声向量映射到高维的数据空间，从而生成新的数据样本。判别器则接收来自生成器的假样本 $G(z)$ 和来自真实数据集的真实样本，并尝试区分它们。判别器的输出是一个概率值，表示输入样本为真实数据的可能性。

在训练过程中，生成器和判别器通过对抗的方式进行优化。生成器尝试生成难以被判别器识别的数据，而判别器则不断提高其能力以正确分类真实和生成的数据，损失函数如下：

$$L_{GAN} = E_{x \sim p(x)}[\log D(x)] + E_{z \sim p(z)}[\log(1 - D(G(z)))]$$

其中， $D(x)$ 是判别器对真实样本 x 的判断， $G(z)$ 是生成器生成的假样本， $p(x)$ 是真实数据的分布，而 $p(z)$ 是噪声向量的分布。

1.4 基于扩散模型的图像生成算法

扩散模型的理论基础可以追溯到20世纪初期非平衡热力学领域，该物理现象启发了机器学习领域中扩散模型的构建。扩散模型作为一种图像生成模型，最早由斯坦福大学的Sohl-Dickstein在2015年提出，而后2020年Ho等^[8]提出的去噪扩散概率模型(denoising diffusion probabilistic models, DDPM)标志着该模型在图像生成技术上取得了重要的进展。

扩散模型的核心机制在于模拟数据从其原始

分布向一个简化分布(如高斯分布)的逐步过渡，在过程中逐步地向数据中添加噪声，直至数据完全变为噪声。随后，模型通过逆扩散过程，即去噪步骤，逐步从噪声状态恢复出有价值的数 据。在正向扩散阶段，数据样本的复杂性逐步降低；而在逆向过程中，模型学习如何从简化的噪声分布中重建原始数据的分布，进而生成与训练集中数据相似的新样本。

在扩散模型中，生成过程被分解为多个小步骤，每一步都对应于向数据中添加一定量的噪声，其公式如下：

$$x_t = \sqrt{\alpha_t} x_0 + \sqrt{1 - \alpha_t} \varepsilon$$

其中， x_t 是第 t 步的扩散状态， x_0 是原始数据点， ε 是从高斯分布中采样的噪声，而 α_t 是一个预定的方差调度，控制每一步噪声的添加量。

逆扩散过程，模型需要学习如何逐步去除噪声并恢复出原始数据。这个过程可以通过最小化去噪损失函数来实现，其公式如下：

$$L = \|x_0 - \hat{x}_T\|^2$$

其中， \hat{x}_T 是通过逆扩散过程生成的样本，而 x_0 是原始的真实数据。通过优化这个损失函数，模型学习如何准确地从噪声中恢复数据。

基于深度学习的图像生成算法主要有以上4类，图2归纳了以这四类模型为基础的变体模型。

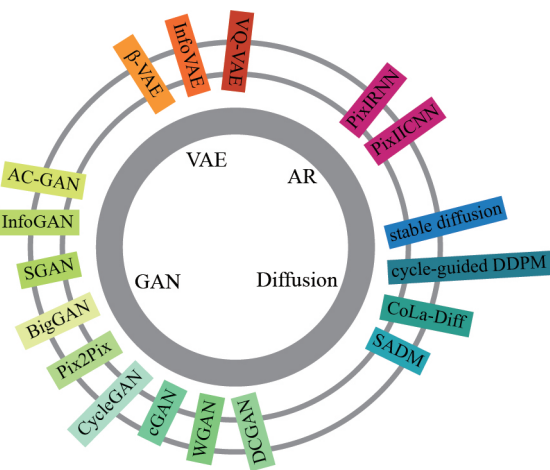


图2 医学影像生成模型概述

Fig. 2 The overview of medical images generation model

2 基于深度学习的医学图像生成

2.1 基于深度学习的CT生成方法

基于深度学习的CT合成技术在捕捉CT与源模态之间复杂的非线性映射方面表现出色，能够生成逼真的合成图像。这些方法在临床和非临床

应用中均已证明其优越性,相比传统方法(如基于图谱和体素的方法)性能显著提升。特别是基于卷积神经网络(convolutional neural network, CNN)和生成对抗网络^[12]的方法在从源图像模态准确估计 CT Hounsfield 单位(HU)方面表现出了巨大潜力^[13]。此外,新型的 Transformer 和扩散模型在生成高保真合成图像方面表现出色^[14]。

深度 3D 网络在编码 MRI 和 CT 之间复杂映射以获得更精确的 HU 预测方面表现更佳。Fu 等^[15]首次展示用于 MRI 到 CT 翻译的端到端 3D CNN 模型,验证了这一点。Zimmermann 等^[16]提出的 3D U-Net 方法则通过多序列 MRI 生成 CT 图像,有效提高了 CT 图像的生成质量。此外,3D cGAN 的方法被广泛用于解决 2D 网络中的图像切片不连续性问题。为了降低 3D 深度网络的高计算成本,部分研究采用创新方法,利用 2D 图像保留图像的 3D 结构信息。Pandeva 和 Schubert^[17]提出了一种引入编码器的生成对抗网络,用于保留多模态分布。

从头颈 MRI 图像预测 CT 图像具有挑战性,因为患者解剖结构和成像协议的多样性。Touati 等^[18]提出了一种新颖的特征不变方法,以匹配合成的和真实 CT 图像在颈部和头部区域的共同结构细节。胸部图像的合成也由于肺的电子密度的异质性和建模病变的难度而具有挑战性^[15]。通过为肿瘤区域分配高密度值的混合方法,研究者成功地展示了合成胸部 CT 成像在准确复制复杂肺区域细节方面的能力。

在 MRI 和 CT 图像之间映射时,许多基于 GAN 的 CT 合成方法依赖于使用单一 MR 序列映射 CT 图像^[19]。然而,这些方法没有解决 MR 对比度和协议的变化问题。增强型 CycleGAN(AugCycleGAN)是一种创新方法,通过使用多中心数据提高 CycleGAN 的普遍性来解决这一限制^[20]。

最近的大多数医学图像翻译方法专注于翻译整个图像的效果,而不是特定的感兴趣区域。Emami 等^[21]提出了一种基于注意力的 GAN 方法,通过在判别器中计算空间注意力来协助生成器关注真实与合成 CT 图像之间差距较大的区域。在合成 CT 图像时保留高频细节是一大挑战。标准损失函数(包括 L1 损失)在低频图像内容中表现良好,但在高频细节方面效果不佳。Zhao 等^[22]提出了一种新颖的方法,通过将对抗框架与非对抗损失独特结合来捕捉图像目标模态中的高频和低频成分。由于其生成高保真图像的潜在能力, Vision

Transformers 和基于扩散的方法已在最近的文献中用于 CT 合成。结合 CNNs 和 Transformers 的混合架构被引入以捕捉 MR 图像中的多层次信息并合成具有改进的强度和结构细节的 CT。见表 3。

2.2 基于深度学习的 MR 生成方法

利用深度学习技术的多模态 MR 图像合成表现出良好的前景。根据当前关于 MR 模态之间映射的最先进方法,MR 合成可分为基于单模态和多模态的深度学习策略。Chartsias 等^[23]提出一种能够通过现有的模态还原缺失的模态的方法,大多数 MR 对比度合成技术依赖于 GAN 进行图像到图像翻译,利用 2D U-Net 生成器和 PatchGAN 判别器^[24]。此外,Transformer 和基于扩散的方法在最新的 MRI 合成研究中备受关注^[25]。

在 MR 合成中,保留 3D 信息对于提高合成图像的精度和真实性非常重要。传统的对抗性方法难以处理高分辨率和 3D 特征。Zhan 等^[26]提出了一种 3D 多尺度补丁方法,通过低分辨率 GAN 和逐级高分辨率 GAN 提升图像细节。Meng 等^[27]则通过从高频细节中提取深层语义信息,并将这些信息与解码器网络中的特征映射结合,来增强网络的特征。对于多对比域数据,许多 MR 合成技术依赖于两个对比的合成,如从 T2 加权(T2W) MRI 生成 T1 加权(T1W) MRI。同时,也有研究探索多模态图像合成,以充分利用多模态 MRI 中获取的解剖特征。Chartsias 等^[23]展示了一种基于模态不变潜在表示的学习方法,将所有输入模态映射到共享的潜在空间中。

虽然多模态 MRI 合成具有从多个图像对比中学习共享特征的优势,尤其是当特征在单个源模态中表示较弱时,Yurt 等^[28]发现完全忽略图像对之间一对一翻译是一个重大问题,因此开发了多流 GAN 模型,利用不同模态的共享和互补图像特征来充分利用单模态和多模态图像合成的优势。与传统的基于对抗的生成方法相比, Vision Transformers 由于其良好的性能和在医学成像任务中建模上下文数据表示的能力而引起了更多的关注。Dalmaz 等^[29]引入了首个用于医学图像合成的基于 Transformer 的生成器模型,称为 Residual Vision Transformers (ResViT)。基于得分的生成模型通过随机扩散技术在高效采样目标分布方面表现出色。Qin 等^[30]通过引入无分类器的条件扩散方法提出了首个基于得分的跨模态 MRI 合成模型。见表 4。

表 3 基于深度学习的CT生成综述
Tab. 3 Reviews on CT generation based on deep learning

生成来源及论文	方法	损失函数	备注
MR→CT			
Nie, et al (2017, 2018) ^[24]	Cascade GAN	L1, 2, 4	脑部, 盆腔
Emami, et al (2018) ^[13]	cGAN	L1, 2	脑部
Liu, et al (2020) ^[34]	2.5D U-Net	加权MSE	腹部
Sun, et al (2022) ^[35]	2.5D CycleGAN	AL + CCL + 形状一致性损失	脑部
Zeng, et al (2019) ^[36]	(2D, 3D) cGAN	AL + L1	头部
Qi, et al (2020) ^[19]	2D cGAN	AL + L1 + Softmax	腹部
Li, et al (2019) ^[37]	2D U-Net	聚焦组织MSE	脑部
Spadea, et al (2019) ^[38]	2D U-Net	MAE	头部
Abu-Srhan, et al (2021) ^[39]	2D CycleGAN	AL + L1	脑部
Kearney, et al (2020) ^[40]	2D CycleGAN	AL + CCL	头部, 颈部, 脑部
Ang, et al (2022) ^[41]	2D cGAN	AL + 风格损失 + 感知损失	头部, 颈部
Shi, et al (2021) ^[42]	3D U-Net	MAE	头部
Touati, et al (2021) ^[18]	2D cGAN	AL + L1	头部, 颈部
Chen, et al (2021) ^[43]	2D CycleGAN	AL + CCL + 交叉熵损失	腹部
Brou Boni, et al (2021) ^[44]	2D CycleGAN	AL + CCL + 边缘匹配损失	膀胱, 直肠, 股骨头
Li, et al (2023) ^[45]	2D CNN	AL + L1	脑部
Tie, et al (2020) ^[46]	2D cGAN	AL + L1	头部
Yang, et al (2018) ^[47]	2D CycleGAN	AL + CCL + 结构一致性损失	脑部
Wang, et al (2021) ^[48]	2D CycleGAN	AL + CCL	盆腔
Li, et al (2021) ^[49]	2.5D CycleGAN	AL + CCL	脑部
Li, et al (2023) ^[50]	2D CNN+ Transformer	AL + L1 + MS-SSIM	头部, 颈部
Zhao, et al (2023) ^[51]	2D CNN+ Transformer	AL + 特征重建损失 + MSE	盆腔
Lyu, et al (2022) ^[52]	2D DDPM	MSE	盆腔
PET→CT			
Armanious, et al (2018) ^[53]	cGAN*	L1, 2, 8, 11	脑部
Dong, et al (2019) ^[54]	3D CycleGAN	AL + CCL + GDL	全身
Armanious, et al (2019) ^[55]	2D CycleGAN	AL + CCL + 风格损失 + 感知损失	脑部

2.3 基于深度学习的CTA生成方法

医学图像生成是一个快速发展的领域，它利用深度学习技术从低剂量或非对比CT扫描中产生高质量的图像。这些图像可以达到与对比增强CTA相当的诊断性能且无需注射造影剂。例如，Chen等^[31]开发了一种级联生成辅助网络，通过CT平扫合成二维CTA切片来诊断主动脉夹层。他们的方法结合了图像配准、nnU-Net分割和带有DCT通道关注机制的双鉴别器网络。Lyu等^[32]提出了一种基于生成式对抗网络的模型，该模型在不使用造影剂的情况下生成颈部和腹部CTA样图像，在外部验证集上实现了0.906的结构相似性。Killekar等^[33]使用条件GAN从完全配准的无对比薄层CT生成伪对比CTA。Lyu等专注于颈部和腹部大血管的生成，Killekar等专注于心脏解剖的评估。他们的方法有可能在某些临床应用中取代造影剂CTA

和伪造影剂CTA，如心脏解剖评估。见表5。

3 讨论

医学影像生成是医学影像研究中一个充满前景的领域，为医学研究提供潜在应用场景^[69-70]，例如通过跨模态转换，可实现无需造影剂的CT/MRI直接生成PET/CTA影像，有望解决造影剂过敏或肾功能不全患者无法做检查问题。本文总结了近期医学影像生成研究的文献，通过介绍4种生成领域的经典模型，总结在不同模态转化情况下，生成模型的适用性和效果及可能促成相应的临床应用。

然而，该领域仍面临若干挑战。(1)模态间差异性：针对不同模态的特性，需优化模型架构和损失函数，以实现更准确的跨模态影像生成；(2)数据配对问题：医学影像的采集成本高昂，获

表 4 基于深度学习的MR生成综述
Tab. 4 Summary of MR generation based on deep learning

生成来源及论文	方法	损失函数	备注
CT→MR			
Jin, et al (2018) ^[56]	CycleGAN	L1, 2, 3	脑部
Jiang, et al (2018) ^[57]	CycleGAN*	L1, 2, 3, 7, 8	肺部
PET→MR			
Choi, et al (2017) ^[58]	pix2pix	L1, 2	脑部
T1↔ T2 MR			
Dar, et al (2019) ^[59]	2D cGAN	AL+ L1+ 感知损失	
Dai, et al (2020) ^[60]	2D cGAN	AL + 分类损失 + CCL + 一致性损失	
Zhao, et al (2021) ^[61]	3D CycleGAN	AL + CCL + 感知损失	
Özbey, et al (2022) ^[25]	基于扩散过程的2D CycleGAN	AL + CCL	
Wang, et al (2023) ^[62]	2D DDPM	MAE	
T1→ FLAIR			
Yu, et al (2018) ^[63]	cGAN	L1, 2	脑部
Zhan, et al (2021) ^[14]	2D cGAN	AL + L1+ GDL	
Yurt, et al (2021) ^[28]	2D cGAN	AL + L1	
FLAIR→ T1			
Li, et al (2019) ^[64]	2D CycleGAN	AL + CCL	
T1, T2→ MRA			
Olut, et al (2018) ^[65]	pix2pix*	L1, 2, 13	脑部
3T→ 7T MR			
Nie, et al (2018) ^[66]	Cascade GAN	L1, 2, 4	脑部
MR↔ CT			
Yang, et al (2018) ^[47]	CycleGAN*	L1, 2, 3, 10	脑部
Maspero, et al (2018) ^[67]	pix2pix	L1, 2	盆腔

表 5 基于深度学习的CTA生成综述
Tab. 5 Summary of CTA generation based on deep learning

生成来源及论文	方法	损失函数	备注
CT→ CTA			
Hua, et al (2024) ^[68]	VesselTransGAN	L1, 2	脑部, 颈部
Chen, et al (2020) ^[31]	nnU-Net+ Cascade GAN	Dice	主动脉
Lyu, et al (2023) ^[32]	GAN	L1	脑部, 腹部
Killekar , et al (2023) ^[33]	cGAN		心脏

取大量高质量的成对数据样本在医学影像领域尤为困难；(3)计算成本问题。大多数研究采用复杂的深度学习架构，而医学影像数据通常是高维的，这就需要强大的计算资源来支持模型的高效训练。

未来的研究工作可以聚焦于以下几个关键领域：提升数据多样性和丰富性，加强跨模态影像

的生成与融合技术，优化实时处理算法。构建规模更大、更具代表性的医学影像数据集，以支持生成模型的训练和优化；探索多模态影像融合技术，以增强临床诊断的准确性；提高数据采集速度和算法效率，以满足临床应用的需求。

本综述详细调研了关于医学影像生成的最新文献，通过模型构建和损失函数公式详细介绍了经典的4种生成算法，并评估了这4类医学影像生成算法在跨模态医学影像生成方面的表现，以生成CT、生成MRI和生成CTA三个维度为例，研究基于不同部位不同模态间转换的深度学习生成方法。此外，我们确定了当前的挑战和潜在的未来研究方向，为其他研究人员提供借鉴，使医学影像生成技术更好地发展和创新。

作者贡献 花芸、康敏诗：论文撰写和修订；刘盼、郭华源、李轶玮：审读和修订；金晓宇：论文撰写；何昆仑：论文审读，监督指导。

利益冲突 所有作者声明无利益冲突。

参考文献

- Rombach R, Blattmann A, Lorenz D, et al. High-resolution image synthesis with latent diffusion models [C] //2022 IEEE/ CVF Conference on Computer Vision and Pattern Recognition (CVPR) . New Orleans: IEEE, 2022: 10674-10685.
- Yoon JS, Zhang CH, Suk HI, et al. SADM: sequence-aware diffusion model for longitudinal medical image generation [C] // Information Processing in Medical Imaging. Cham: Springer Nature Switzerland, 2023: 388-400.
- Pan SY, Chang CW, Peng JB, et al. Cycle-guided denoising diffusion probability model for 3D cross-modality MRI synthesis [EB/OL] . <https://arxiv.org/abs/2305.00042v1>.
- Jiang L, Mao Y, Wang XF, et al. CoLa-diff: conditional latent diffusion model for multi-modal MRI synthesis [C] //Medical Image Computing and Computer Assisted Intervention MICCAI 2023. ACM, 2023: 398-408.
- Van Den Oord A, Kalchbrenner N, Vinyals O, et al. Conditional image generation with PixelCNN decoders [C] // Proceedings of the 30th International Conference on Neural Information Processing Systems. Barcelona: ACM, 2016: 4797-4805.
- Kingma DP, Welling M. Auto-encoding variational Bayes [EB/OL] . <https://arxiv.org/abs/1312.6114v11>.
- Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial networks [J] . Commun ACM, 2020, 63 (11) : 139-144.
- Ho J, Jain A, Abbeel P, et al. Denoising diffusion probabilistic models [C] //Proceedings of the 34th International Conference on Neural Information Processing Systems. Vancouver: ACM, 2020: 6840-6851.
- Saharia C, Chan W, Saxena S, et al. Photorealistic text-to-image diffusion models with deep language understanding [EB/OL] . <https://arxiv.org/abs/2205.11487v1>.
- Ramesh A, Dhariwal P, Nichol A, et al. Hierarchical text-conditional image generation with CLIP latents [EB/OL] . <https://doi.org/10.48550/arXiv.2204.06125>.

- 11 Wang SH, Sun Y, Xiang Y, et al. ERNIE 3.0 titan: exploring larger-scale knowledge enhanced pre-training for language understanding and generation [EB/OL]. <https://arxiv.org/abs/2112.12731v1>.
- 12 Singh N, Raza K. Medical image generation using generative adversarial networks: a review [M] //Patgiri R, Biswas A, Roy P. Health Informatics: A Computational Perspective in Healthcare. Springer, 2021.
- 13 Emami H, Dong M, Nejad-Davarani SP, et al. Generating synthetic CTs from magnetic resonance images using generative adversarial networks [J/OL]. <https://doi.org/10.1002/mp.13047>.
- 14 Zhan B, Li D, Wang Y, et al. LR-cGAN: Latent representation based conditional generative adversarial network for multi-modality MRI synthesis [J]. Biomed Signal Process Contr, 2021, 66: 102457.
- 15 Fu YB, Lei Y, Wang TH, et al. Deep learning in medical image registration: a review [J]. Phys Med Biol, 2020, 65 (20): 20TR01.
- 16 Zimmermann L, Knäusel B, Stock M, et al. An MRI sequence independent convolutional neural network for synthetic head CT generation in proton therapy [J]. Z Med Phys, 2022, 32 (2): 218-227.
- 17 Pandeva T, Schubert M. MMGAN: generative adversarial networks for multi-modal distributions [J/OL]. <https://doi.org/10.48550/arXiv.1911.06663>.
- 18 Touati R, Le WT, Kadoury S. A feature invariant generative adversarial network for head and neck MRI/CT image synthesis [J]. Phys Med Biol, 2021, 66 (9): 095001.
- 19 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.
- 20 Li JJ, Chen EP, Ding ZM, et al. Cycle-consistent conditional adversarial transfer networks [C] //Proceedings of the 27th ACM International Conference on Multimedia. ACM, 2019: 747-755.
- 21 Emami H, Dong M, Glide-Hurst CK. Attention-guided generative adversarial network to address atypical anatomy in synthetic CT generation [C] //2020 IEEE 21st International Conference on Information Reuse and Integration for Data Science (IRI). IEEE, 2020: 188-193.
- 22 Zhao YH, Wu RH, Dong H. Unpaired image-to-image translation using adversarial consistency loss [J/OL]. <https://doi.org/10.48550/arXiv.2003.04858>.
- 23 Chartsias A, Joyce T, Giuffrida MV, et al. Multimodal MR synthesis via modality-invariant latent representation [J]. IEEE Trans Med Imaging, 2018, 37 (3): 803-814.
- 24 Nie D, Trullo R, Lian J, et al. Medical image synthesis with context-aware generative adversarial networks [J]. Med Image Comput Comput Assist Interv, 2017, 10435: 417-425.
- 25 Özbey M, Dalmaz O, Dar SUH, et al. Unsupervised medical image translation with adversarial diffusion models [J]. IEEE Trans Med Imag, 2023, 42 (12): 3524-3539.
- 26 Zhan B, Li D, Wu X, et al. Multi-modal MRI image synthesis via GAN with multi-scale gate mergence [J]. IEEE J Biomed Health Inform, 2022, 26 (1): 17-26.
- 27 Meng XX, Sun KC, Xu J, et al. Multi-modal modality-masked diffusion network for brain MRI synthesis with random modality missing [J]. IEEE Trans Med Imag, 2024, 43 (7): 2587-2598.
- 28 Yurt M, Dar SU, Erdem A, et al. mustGAN: multi-stream generative adversarial networks for MR image synthesis [J]. Med Image Anal, 2021, 70: 101944.
- 29 Dalmaz O, Yurt M, Cukur T. ResViT: residual vision transformers for multimodal medical image synthesis [J]. IEEE Trans Med Imaging, 2022, 41 (10): 2598-2614.
- 30 Qin ZW, Liu Z, Zhu P, et al. Style transfer in conditional GANs for cross-modality synthesis of brain magnetic resonance images [J]. Comput Biol Med, 2022, 148: 105928.
- 31 Chen HW, Yan SN, Xie MX, et al. Application of cascaded GAN based on CT scan in the diagnosis of aortic dissection [J]. Comput Methods Programs Biomed, 2022, 226: 107130.
- 32 Lyu JH, Fu Y, Yang ML, et al. Generative adversarial network-based noncontrast CT angiography for aorta and carotid arteries [J]. Radiology, 2023, 309 (2): e230681.
- 33 Killekar A, Kwiecinski J, Kruk M, et al. Pseudo-contrast cardiac CT angiography derived from non-contrast CT using conditional generative adversarial networks [J/OL]. <https://doi.org/10.1117/12.2654592>.
- 34 Liu LL, Johansson A, Cao Y, et al. Abdominal synthetic CT generation from MR Dixon images using a U-Net trained with 'semi-synthetic' CT data [J]. Phys Med Biol, 2020, 65 (12): 125001.
- 35 Sun B, Jia SF, Jiang XL, et al. Double U-Net CycleGAN for 3D MR to CT image synthesis [J]. Int J Comput Assist Radiol Surg, 2023, 18 (1): 149-156.
- 36 Zeng GD, Zheng GY. Hybrid generative adversarial networks for deep MR to CT synthesis using unpaired data [C] //Medical Image Computing and Computer Assisted Intervention MICCAI 2019. Cham: Springer International Publishing, 2019: 759-767.
- 37 Li YF, Li W, He P, et al. CT synthesis from MRI images based on deep learning methods for MRI-only radiotherapy [C] //2019 International Conference on Medical Imaging Physics and Engineering (ICMIPE). IEEE, 2019: 1-6.
- 38 Spadea MF, Pileggi G, Zaffino P, et al. Deep convolution neural network (DCNN) multiplane approach to synthetic CT generation from MR images-application in brain proton therapy [J]. Int J Radiat Oncol Biol Phys, 2019, 105 (3): 495-503.
- 39 Abu-Srhan A, Almallahi I, Abushariah MAM, et al. Paired-unpaired Unsupervised Attention Guided GAN with transfer learning for bidirectional brain MR-CT synthesis [J]. Comput Biol Med, 2021, 136: 104763.
- 40 Kearney V, Ziemer BP, Perry A, et al. Attention-aware discrimination for MR-to-CT image translation using cycle-consistent generative adversarial networks [J]. Radiol Artif Intell, 2020, 2 (2): e190027.
- 41 Ang SP, Lam Phung S, Field M, et al. An improved deep learning framework for MR-to-CT image synthesis with a new hybrid objective function [C] //2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI). IEEE, 2022: 1-5.
- 42 Shi ZL, Mettes P, Zheng GY, et al. Frequency-supervised MR-to-CT image synthesis [C] //Deep Generative Models, and Data Augmentation, Labelling, and Imperfections. Cham: Springer International Publishing, 2021: 3-13.
- 43 Chen JX, Wei J, Li R. TarGAN: target-aware generative adversarial networks for multi-modality medical image translation [C] //Medical Image Computing and Computer Assisted Intervention MICCAI 2021. Cham: Springer International Publishing, 2021: 24-33.
- 44 Brou Boni KND, Klein J, Gulyban A, et al. Improving generalization in MR-to-CT synthesis in radiotherapy by using an augmented cycle generative adversarial network with unpaired data [J]. Med Phys, 2021, 48 (6): 3003-3010.
- 45 Li Y, Xu SS, Chen HB, et al. CT synthesis from multi-sequence MRI using adaptive fusion network [J]. Comput Biol

- Med, 2023, 157: 106738.
- 46 Tie X, Lam SK, Zhang Y, et al. Pseudo-CT generation from multi-parametric MRI using a novel multi-channel multi-path conditional generative adversarial network for nasopharyngeal carcinoma patients [J]. Med Phys, 2020, 47 (4): 1750-1762.
- 47 Yang HR, Sun J, Carass A, et al. Unpaired brain MR-to-CT synthesis using a structure-constrained CycleGAN [C] //Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Cham: Springer International Publishing, 2018; 174-182.
- 48 Wang C, Uh J, He XX, et al. Transfer learning-based synthetic CT generation for MR-only proton therapy planning in children with pelvic sarcomas [C] //Medical Imaging 2021: Physics of Medical Imaging. SPIE, 2021; 150.
- 49 Li W, Kazemifar S, Bai T, et al. Synthesizing CT images from MR images with deep learning: model generalization for different datasets through transfer learning [J]. Biomed Phys Eng Express, 2021, 7 (2): 025020.
- 50 Li Y, Xu SS, Lu Y, et al. CT synthesis from MRI with an improved multi-scale learning network [J]. Front Phys, 2023, 11: 1088899.
- 51 Zhao B, Cheng TT, Zhang XR, et al. CT synthesis from MR in the pelvic area using Residual Transformer Conditional GAN [J]. Comput Med Imaging Graph, 2023, 103: 102150.
- 52 Lyu Q, Wang G. Conversion between CT and MRI images using diffusion and score-matching models [EB/OL]. <https://arxiv.org/abs/2209.12104v2>.
- 53 Armanious K, Jiang CM, Fischer M, et al. MedGAN: Medical image translation using GANs [J]. Comput Med Imag Graph, 2020, 79: 101684.
- 54 Dong X, Wang TH, Lei Y, et al. Synthetic CT generation from non-attenuation corrected PET images for whole-body PET imaging [J]. Phys Med Biol, 2019, 64 (21): 215016.
- 55 Armanious K, Jiang CM, Abdulatif S, et al. Unsupervised medical image translation using cycle-MedGAN [C] //2019 27th European Signal Processing Conference (EUSIPCO). IEEE, 2019: 1-5.
- 56 Jin CB, Kim H, Liu MJ, et al. Deep CT to MR synthesis using paired and unpaired data [J]. Sensors (Basel), 2019, 19 (10): 2361.
- 57 Jiang J, Hu YC, Tyagi N, et al. Tumor-aware, adversarial domain adaptation from CT to MRI for lung cancer segmentation [J]. Med Image Comput Comput Assist Interv, 2018, 11071: 777-785.
- 58 Choi H, Lee DS, Alzheimer's Disease Neuroimaging Initiative. Generation of structural MR images from amyloid PET: application to MR-less quantification [J]. J Nucl Med, 2018, 59 (7): 1111-1117.
- 59 Dar SU, 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.
- 60 Dai XJ, Lei Y, Fu YB, et al. Multimodal MRI synthesis using unified generative adversarial networks [J]. Med Phys, 2020, 47 (12): 6343-6354.
- 61 Zhao P, Pan H, Xia SY. MRI-trans-GAN: 3D MRI cross-modality translation [C] //2021 40th Chinese Control Conference (CCC). IEEE, 2021; 7229-7234.
- 62 Wang ZH, Yang YY, Sermesant M, et al. Zero-shot-learning cross-modality data translation through mutual information guided stochastic diffusion [EB/OL]. <https://doi.org/10.48550/arXiv.2301.13743>.
- 63 Yu BT, Zhou LP, 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). IEEE, 2018: 626-630.
- 64 Li HW, Paetzold JC, Sekuboyina A, et al. DiamondGAN: unified multi-modal generative adversarial networks for MRI sequences synthesis [C] //Medical Image Computing and Computer Assisted Intervention MICCAI 2019. Cham: Springer International Publishing, 2019: 795-803.
- 65 Olut S, Sahin YH, Demir U, et al. Generative adversarial training for MRA image synthesis using multi-contrast MRI [C] //Predictive Intelligence in Medicine. Cham: Springer International Publishing, 2018: 147-154.
- 66 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.
- 67 Maspero M, Savenije MHF, Dinkla AM, et al. Dose evaluation of fast synthetic-CT generation using a generative adversarial network for general pelvis MR-only radiotherapy [J]. Phys Med Biol, 2018, 63 (18): 185001.
- 68 Hua Y, Bian SY, Liu P, et al. VesselTransGAN to CT imaging: a contrast medium free CTA solution [J]. IEEE Access, 2024, 12: 129917-129926.
- 69 Dayarathna S, Islam KT, Uribe S, et al. Deep learning based synthesis of MRI, CT and PET: review and analysis [J]. Med Image Anal, 2024, 92: 103046.
- 70 Yi X, Walia E, Babyn P. Generative adversarial network in medical imaging: a review [J]. Med Image Anal, 2019, 58: 101552.

(责任编辑:孟晓彤)