

基于深度学习的医学图像合成：方法、应用与未来方向

杨莉

(重庆医科大学附属康复医院, 重庆 400013)

摘要: 随着医学影像技术的快速发展, 深度学习在医学图像合成领域展现出巨大潜力, 为临床诊断、治疗规划和医学研究提供了新机遇。本文系统综述了基于深度学习的医学图像合成方法, 重点探讨了生成对抗网络 (GANs) 在跨模态合成和域适应中的应用, 扩散模型 (DM) 的高保真度生成能力及其在多模态合成中的局限性 (如合成样本多样性不足), 以及 Transformer 模型和混合模型与领域适应技术的优势。在应用场景方面, 本文分析了跨模态图像合成 (例如 MRI 到 CT、MRI 到 PET 转换)、模态内图像增强与标准化、图像重建与去噪、纵向图像生成与疾病进展建模、特定病变或组织合成等实践案例, 同时指出现有方法在泛化性、计算效率和伦理合规性方面的局限。未来研究应聚焦于提升模型的可解释性与可靠性、多模态融合能力及临床应用可行性。

关键词: 深度学习; 医学图像合成; 生成对抗网络; 扩散模型; 跨模态合成

DOI: doi.org/10.70693/cjmsr.v1i3.1764

医学图像合成作为人工智能与医学影像学交叉的重要研究方向, 旨在利用算法生成或重建高质量的医学影像, 以弥补数据采集中的不足并提升临床诊断的精度。近年来, 随着计算能力的提升与海量医学影像数据的积累, 深度学习技术在该领域展现出强大的潜力。尤其是卷积神经网络、生成对抗网络 (Generative Adversarial Networks, GANs) 以及扩散模型 (Diffusion Models, DM) 等方法的快速发展, 使得医学图像合成在跨模态转换、图像重建、数据增强及隐私保护等方面取得显著进展。

在临床应用中, 深度学习驱动的合成技术为解决多项关键问题提供了新思路。例如, 通过低剂量计算机断层扫描 (Computed Tomography, CT) 生成高质量图像以减少辐射暴露; 利用磁共振成像 (Magnetic Resonance Imaging, MRI) 模态合成弥补数据缺失以辅助诊断; 通过多模态融合提升病灶检测的准确性等。这些进展不仅改善了医学影像质量, 也促进了影像组学与精准医疗的发展。然而, 尽管成果丰富, 但当下研究仍面临模型可解释性不足、数据分布偏差与临床可推广性等挑战。

因此, 本文旨在系统梳理深度学习在医学图像合成中的主要方法, 应用进展及发展趋势。分析不同模型的技术特点与适用场景, 并探讨该领域的关键挑战及未来发展方向。通过全面总结现有研究成果, 本文期望为医学图像智能生成的理论与临床发展提供思路。

1 深度学习医学图像合成的主要方法

医学图像合成在临床诊断与治疗中具有重要意义, 其发展得益于深度学习技术的进步。本章旨在全面回顾深度学习在医学图像合成中的核心技术, 重点探讨 GANs、DM、Transformer 模型及其混合架构与领域适应技术。

1.1 基于 GANs 的医学图像跨模态合成和域适应

GANs 是一种深度学习框架, 通过对抗训练的方式生成高质量的图像。在医学图像跨模态合成任务中, GANs 通过学习目标模态图像的生成分布来生成与输入模态图像具有相似结构和特征的输出图像^[1]。多样化的 GAN 变体在跨模态合成和域适应中展现了架构创新。循环一致性生成对抗网络 (Cycle-Consistent GAN, CycleGAN) 是一种无监督学习方法, 特别适用于没有配对数据的情况下进行跨模态合成。对于医学图像来说, CycleGAN 能够实现 CT 和 MRI 之间的图像转换, 而无需成对的训练数据^[2]。其关键优势在于, CycleGAN 通过引入循环一致性^[3,4]损失, 使得生成的图像能够在两个模态间相互转换而保持一致性。在 3D 图像合成方面, 3D 可逆 GAN 仅需一个生成器即可实现 MRI 与正电

作者简介: 杨莉(1993—), 女, 生物医学工程专业, 从事放射医学技术

子发射型计算机断层显像 (Positron Emission Computed Tomography, PET) 的双向合成, 具有计算效率优势^[5]。

1.2 DM 在医学图像生成中的高保真度与多模态合成能力

DM 通过多次噪声添加将原始图像转化为随机噪声, 然后通过训练得到的反向过程将噪声逐步去除, 最终生成逼真的图像。其核心优势是高保真度和多模态合成能力。其中, 高保真度体现在生成图像的细节和精确度上, 而多模态合成能力则指的是通过同一个模型, 能够实现不同类型 (如 CT、MRI、超声等) 医学图像的生成与转换。去噪扩散概率模型, 通过模拟逐步加噪的前向扩散过程和学习逆向去噪过程来生成高质量图像, 其灵感来源于非平衡热力学^[6]。这种渐进式的去噪机制使得扩散模型能够生成具有高保真度和多样性的图像。此外, 联合扩散注意力模型通过学习 MRI 和噪声化 PET 图像的联合概率分布, 并利用预测-校正器机制进行采样, 生成高保真度的 PET 图像, 弥补了传统扩散模型在条件生成上的不足^[7]。

1.3 Transformer 模型

Transformer 模型凭借其全局注意力机制, 在处理复杂解剖结构和维持图像全局一致性方面展现出优于传统卷积神经网络的潜力。其能够捕捉长距离依赖关系, 有助于在医学图像合成中提高对非典型解剖结构和病变的泛化能力^[8]。在医学图像合成领域, 多尺度 Transformer 网络被提出用于跨模态 MR 图像合成^[9]。在 PET 图像合成方面, 医学图像合成领域的一种深度学习模型 (Frequency-domain Spatial Attention Network, FrSA-Net, FrSA-Net) 将 Transformer 与频率域注意力机制相结合, 以解决 PET 图像固有的低信噪比和稀疏性问题, 并在脱髓鞘预测中展现潜力^[10]。空间自适应和 Transformer 融合网络利用 Transformer 融合编码器捕捉 PET 和 MRI 之间的局部和全局关系, 以实现高层特征融合^[11]。

1.4 混合模型与领域适应技术

混合模型指的是将多个深度学习模型的优点结合在一起, 形成一个新的复合模型。通过融合不同模型的优势, 混合模型能够弥补单一模型的局限性, 提高图像合成的效果和泛化能力。混合模型包括: 结合 VAE 和 GAN 技术用于 tagged MRI 到 cine MRI 转换的模型^[12]; Hi-Net 融合模态特定编码器-解码器与多模态融合模块; 以及 MMTrans 模型^[13,14]结合 Swin Transformer 与 U-Net 结构的形变注册网络。此外, ResViT 模型^[8]将 Transformer 模块与 CNN 模块相结合, 用于医学图像合成。

领域适应是指将一个领域 (源领域) 上的学习到的知识迁移到另一个领域 (目标领域) 上, 以弥补目标领域数据不足的问题。在医学图像中, 由于数据的标注困难和不同医院、设备、协议等因素的影响, 同一类型的医学图像往往会存在较大差异。因此, 领域适应技术在医学图像合成中具有重要作用, 能够提高跨不同设备、不同医院和不同病种之间图像合成的效果。CycleGAN 框架^[15,16]作为一种非配对图像转换模型, 本身也具有一定的域适应能力。迁移学习^[15]在 CycleGAN 模型上的应用, 通过预训练和微调提升模型在不同数据集上的泛化能力, 也属于域适应的一种形式。

2 深度学习医学图像合成的应用场景

深度学习在医学图像合成领域展现出广泛的应用前景, 这些应用共同推动了医学影像技术的进步, 提升了诊断和治疗的精准性与效率。应用场景主要体现在以下几个方面:

2.1 跨模态医学图像合成: MRI 到 CT、MRI 到 PET 等

跨模态医学图像合成在放射治疗计划和诊断辅助方面表现突出, 在 MRI 到 CT 合成任务中, 模型性能的评估主要集中在骨骼结构还原、软组织细节保留、局部目标区域质量以及体积一致性方面。多项研究致力于利用 MRI 数据生成合成 CT (sCT) 图像, 以减少患者的辐射暴露并弥补 MRI 在电子密度信息上的不足, 例如脑肿瘤放射治疗、肺部 MR 引导放疗、前列腺癌 MR 引导自适应放疗以及头颈部区域的仅使用 MRI 的放疗^[16,17]。

在 MRI 到 PET 合成方面, shin^[18]的研究独具特色, 其核心应用在于利用 T1 加权 MRI 图像生成 PET 图像, 并将其恢复到原始的标准摄取值比范围, 有效克服了直接从 MRI 图像预测放射性示踪剂摄取的挑战。此外, 其他研究也探索了利用 MRI 信息合成 PET 图像, 例如基于 Transformer 和频率域注意力机制的 FrSA-Net 模型旨在提高 PET 图像质量, 辅助诊断多发性硬化症^[10]。谢等人^[7]的研究利用高

场和超高场 MRI 合成 PET 图像,并捕捉生理分布和真实解剖结构。通过 MRI 数据重构缺失的 PET 信息,克服了 PET 扫描的限制^[5]。

2.2 模态内图像增强与域内标准化

在图像增强方面,研究人员利用深度学习方法提升低场强 MRI 的图像质量。例如,通过合成技术将低场强 MRI 图像转化为高质量的高场强 MRI 图像,从而克服低场强图像噪声大、分辨率低、对比度差等固有缺陷,显著提高诊断精度^[19]。例如,3T MRI 到 7T MRI 的合成也被视为模态内图像质量增强的应用。除了 MRI,深度学习在 PET 图像增强方面也取得显著进展。多项研究致力于将低剂量或低计数 PET 图像恢复至接近全剂量或标准剂量的质量水平,以提高图像的定量准确性和诊断实用性^[10,11]。

在域内标准化方面,深度学习能够统一不同对比度或采集参数的 MRI 图像,提高诊断的一致性和可靠性。例如,U-Net 模型能够将一种 MRI 对比度(如 T1、T2 或 FLAIR)转换为另一种对比度,生成缺失的 MRI 序列以填补扫描空白^[8,20]。通过合成缺失的序列,可以为下游任务提供更完整的数据。此外,深度学习方法也被用于处理运动伪影、弥补低分辨率 MRI 的细节不足,如将低分辨率 tagged MRI 合成高分辨率 cine MRI^[3,12]。在某些情况下,通过调整标签生成具有不同特征(如肿瘤大小、位置)的合成 MRI,也能视为一种模态内的增强或域内标准化^[20]。

2.3 图像重建、去噪与逆问题学习

为重建目标图像从欠采样或噪声数据中恢复高质量医学图像方面,深度学习方法被广泛应用。例如,多项研究聚焦于从低剂量或低计数 PET 图像中估计全剂量 PET 图像,以提升图像质量并降低患者的辐射暴露^[11]。通过学习低剂量 PET 和 MRI 之间的映射关系,可以有效地解决低剂量 PET 图像的噪声问题。

扩散模型固有的去噪过程可视为求解逆问题,能够从噪声中逐步恢复数据^[6]。特别是针对稀疏视图测量嵌入的应用,DDMM-Synth 通过结合扩散模型和测量嵌入的反向采样,将 CT 重建视为一个逆问题,并利用 MRI 作为引导,同时处理带有噪声的测量,间接实现了去噪。Wang 等人^[9]提出的通用零样本框架 DDNM 和 DDNM+,则通过在反向扩散过程中优化零空间内容,解决了多种线性图像复原问题,包括去噪和修复等。这些进展共同推动了医学图像高质量重建和去噪技术的发展。

在解决医学图像逆问题中,Schwab 等人^[21]提出了“零空间网络”的概念,通过结合经典重建方法和神经网络来优化病态逆问题,并通过保证数据一致性来提升重建精度。此外,DeepPET 模型将 PET 正弦图数据作为输入,通过深度卷积编码器-解码器网络直接输出重建图像,绕过了传统迭代算法,显著提高了重建速度和图像质量^[22]。

2.4 纵向图像生成与疾病进展建模

深度学习方法可捕捉图像随时间变化的动态特征,尤其对于纵向图像生成与疾病进展建模具有重要意义。序列感知扩散模型(Sequence-Aware Diffusion Model,SADM)是一种专门用于生成纵向医学图像的方法^[23]。SADM 的核心在于引入了基于 Transformer 的注意力模块,用以处理序列输入并作为扩散模型的条件模块,从而能够学习纵向依赖性。在推理阶段,SADM 采用自回归生成方式生成一系列图像,此类模型在模拟疾病进展、预测治疗反应和评估长期疗效方面具有巨大潜力。此外,PTNet3D 模型已展示出在处理纵向婴儿脑 MRI 方面的能力,能够合成不同年龄段婴儿的大脑 MRI,即便在传统 CNN 模型难以处理的 0-6 个月婴儿数据上也能取得良好结果^[24]。这对于神经发育轨迹和疾病进展的研究至关重要。还有文献指出,通过合成 PET 图像来捕捉阿尔茨海默病患者额叶和顶叶的低代谢区域,有助于疾病进展的分析和建模^[7]。

然而,需要指出的是,当前大量研究仍主要集中于单次扫描或单个时间点的图像合成,例如 MR 到 CT 合成、单次 MRI 扫描的对比度合成或静态 PET 图像合成,尚未直接涉及纵向图像生成或疾病进展建模^[2,5,10,11-16,22]。未来研究应着重于进一步开发和优化序列感知模型,以更准确地捕捉疾病进展的复杂动态,并将其应用于临床决策支持和个性化治疗方案的制定。

2.5 特定病变或组织合成

在医学图像合成领域,针对特定病变或组织的合成相较于整体图像合成,面临独特挑战,并侧重

于真实性与诊断可用性。此类方法旨在生成具有特定病理特征或解剖结构的图像，以辅助诊断、治疗规划及研究。例如，Wallrodt 等人旨在生成带有特定病变（如溃疡）的内窥镜图像，以提升训练模拟器的真实感。在 PET 图像合成中，FrSA-Net 模型能够更准确地生成预测脱髓鞘的 PET 图像，从而辅助 MS 的诊断和研究^[10]。一些研究在生成整体图像的同时，也强调了对特定解剖结构或病理区域的关注。例如，在合成颅骨伪 CT 时，NetBA 网络被引导关注骨骼、颅缝和空气区域，以提高这些关键结构的合成质量^[24]。Rajagopal 等人能够在合成 PET 图像中插入合成病变，用于评估 MRAC 方法在不同。此外，针对脑肿瘤的合成研究，通过使用 BRATS 数据集中的肿瘤分割标签，调整肿瘤大小和位置，生成带有特定脑肿瘤的合成 MRI 图像，以增强数据集并支持肿瘤分割任务^[14,20]。

2.6 现有方法的局限性

当前深度学习在医学图像合成仍面临多重挑战，GANs 的训练不稳定性和模式崩溃问题仍普遍存在^[7,12,19]。CycleGAN 等方法虽然能够处理未配对数据，但其循环一致性是近似的，需要额外的损失项来强制执行。传统的 GAN 方法在处理 PET 图像的宽强度范围时效果不佳，且可能生成不符合物理规律的伪影或改变关键特征^[16,18]。DM 尽管在图像质量上表现出色，但其效率问题突出，采样速度慢，限制了在需要快速生成图像的临床应用中的可行性。扩散去噪过程可能改变目标图像的原始分布结构，导致忽略解剖结构的一致性，生成临床相关性较低的结果^[12]。Transformer 模型在全局一致性和局部细节之间存在权衡。虽然 Transformer 能够捕捉长距离依赖性并提升全局特征，但在处理大尺寸 3D 医学图像时，面临计算量大、内存消耗高的问题^[13]。2D 合成方法普遍存在 3D 体积不连续性问题，而 3D 方法又面临计算成本高和内存需求大的挑战。此外，模型的泛化能力受限于训练数据集的规模和多样性，特别是在处理异常解剖结构或罕见病灶时表现不佳。

3 未来研究方向

深度学习的快速发展极大地推动了医学图像合成技术的进步但要实现从实验研究到临床实践的全面落地，仍存在诸多挑战与发展空间。未来的研究方向可从以下几个方面进行拓展和深化。

模型的可解释性与可靠性提升：当前多数基于深度学习的生成模型（如 GANs、Diffusion Models）仍属于“黑箱式”结构，缺乏对内部特征提取和生成机制的可解释性。通过可视化特征图来揭示模型学习过程，或生成不确定性图来为模型输出提供初步的可解释性^[24]。注意力机制，尤其是在 Transformer 模型中，被认为是增强可解释性的有效途径，能够展示模型在 MRI 和 CT/PET 之间映射时的关注区域^[8,9,10,18]。一些研究通过评估输出的物理意义，如空气和骨骼区域以及质子射程偏移，间接体现了模型的“可解释性”。在可靠性方面，已有研究尝试通过调整输入标签（如 MMSE 分数或肿瘤大小、位置）来控制生成图像的特定特征^[20]。“零空间网络”则通过确保数据一致性来提高模型可靠性，这可视作为一种可控性体现^[21]。未来研究应着重提高模型的可解释性与可靠性，通过可视化方法、注意力机制以及因果推理模型揭示生成过程中的关键特征，以增强临床可接受度和信任度。同时，应建立统一的模型验证与置信度评估体系，确保合成图像在临床诊断中的安全性与一致性。

多模态与跨域合成的融合发展：医学影像涵盖 CT、MRI、PET、超声等多种模态，不同模态间包含互补信息。研究表明，整合来自同一患者的多种模态信息（如 MRI、CT、PET）能够生成更全面、互补的合成图像，显著提高合成性能。未来的工作可以探索更复杂的融合策略，如多模态注意力机制或图神经网络，并将 Transformer 架构扩展到多模态输入。实现从单一模态向多模态协同生成的转变，这不仅有助于提高图像合成的结构与纹理真实性，还将为复杂疾病的综合诊断和功能影像分析提供更全面的数据支持。

数据高效与隐私保护的学习框架：医学影像数据的隐私性与采集成本限制了模型的训练规模。未来应探索更高效、安全的数据利用机制，如联邦学习、小样本学习和迁移学习方法，以减少对大规模标注数据的依赖。同时，应强化数据匿名化与隐私保护机制，平衡算法性能与伦理安全要求。

生成模型的真实性和评估与标准化：目前，医学图像合成的质量评估尚缺乏统一标准。未来研究可综合采用客观指标（如 FID、SSIM、PSNR 等）与主观专家评估，构建多层次的图像质量评价体系。同时，应推动模型训练、数据预处理及评估流程的标准化，确保实验结果的可复现性与跨机构可比性，为临床推广提供可靠依据。

面向临床的可落地应用与集成系统：未来研究的重点应逐步从算法性能优化转向临床可用性验证。应探索与影像采集设备、PACS 系统及影像组学分析平台的深度集成，构建端到端的智能影像生成与分析系统。此外，通过与医生反馈闭环相结合，可实现模型在实际诊断场景中的持续优化与自适应更新，促进智能合成影像在临床中的规范化应用。

新兴技术与跨学科研究：除了单一模型的创新，混合模型和多模态融合技术也备受关注。Transformer 与 CNN 的混合架构已在医学图像合成中展现潜力，而将扩散模型与对抗性学习相结合的 SynDiff 模型也代表了新兴技术方向^[1,8]。跨学科研究被认为是解决当前挑战和开拓新应用领域的关键。医学影像学、生物医学工程、放射肿瘤学、医学物理学乃至计算机视觉等领域的紧密合作，有助于将理论模型转化为实际的临床应用。例如，与临床专家合作可以更好地理解临床需求，指导模型开发和评估，确保合成图像的临床效用和安全性^[24]。同时，因果推断和神经符号 AI 等新兴计算范式，有望增强模型的解释性和可信度，为医学图像合成带来新的突破^[25]。

总体而言，医学图像合成正经历从“算法驱动”向“临床赋能”的关键转型阶段。未来的研究应在模型可解释性、多模态融合、数据安全、标准化评估及临床集成等方面持续深化，以推动深度学习技术真正服务于精准医疗与智能诊断。随着技术与医学的不断融合，医学图像合成有望成为人工智能推动医学影像创新与变革的重要引擎。

4 总结

深度学习在医学图像合成领域取得了显著进展，特别是在提升图像质量、克服数据稀缺和减少辐射暴露方面展现了巨大潜力。GANs、Transformer 架构和扩散模型等技术在多种任务中表现出色，显著提高了临床诊断和治疗规划的精确性^[4,9,20]。这些技术在 MRI 的放射治疗、数据增强和辅助疾病诊断方面已展现出初步应用价值。尽管成就显著，该领域仍面临泛化能力不足、对大规模配对数据依赖以及 3D 合成内存效率和模型可解释性等挑战^[2,15]。未来的研究应聚焦于提升模型效率、增强泛化能力、发展可解释性模型、融合多模态信息以及完善临床评估体系，以期推动深度学习在医学图像合成领域的临床转化。

参考文献：

- [1] Ozbey M, Dalmaz O, Dar SUH, et al. Unsupervised Medical Image Translation With Adversarial Diffusion Models[J]. IEEE Trans Med Imaging, 2023;42(12):3524-3539.
- [2] 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: 3627-3636.
- [3] Ozonoff A, Schaenman J, Jayavelu ND, et al. Phenotypes of disease severity in a cohort of hospitalized COVID-19 patients: Results from the IMPACC study[J]. EBioMedicine. 2022;83:104208.
- [4] Guogang Cao, Shunkun Liu, Hongdong Mao, et al. Improved CycleGAN for MR to CT synthesis[C/OL]//2021 6th International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS)[J]. IEEE: 2021;205-208.
- [5] Lin W, Lin W, Chen G, et al. Alzheimer's Disease Neuroimaging Initiative. Bidirectional Mapping of Brain MRI and PET With 3D Reversible GAN for the Diagnosis of Alzheimer's Disease[J]. Front Neurosci. 2021;15:646013.
- [6] Vasylechko S, Afacan O, Kurugol S. Self Supervised Denoising Diffusion Probabilistic Models for Abdominal DW-MRI[J]. Comput Diffus MRI. 2023;14328:80-91.
- [7] Xie T, Cao C, Cui ZX, et al. Synthesizing PET images from high-field and ultra-high-field MR images using joint diffusion attention model[J]. Med Phys. 2024;51(8):5250-5269.
- [8] 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.
- [9] Li Y, Zhou T, He K, et al. Multi-Scale Transformer Network With Edge-Aware Pre-Training for Cross-Modality MR Image Synthesis[J]. IEEE Trans Med Imaging. 2023;42(11):3395-3407.

- [10] Wei W, Poirion E, Bodini B, et al. Predicting PET-derived demyelination from multimodal MRI using sketcher-refiner adversarial training for multiple sclerosis[J]. *Med Image Anal.* 2019;58:101546.
- [11] Zhang L, Xiao Z, Zhou C, et al. Spatial adaptive and transformer fusion network (STFNet) for low-count PET blind denoising with MRI[J]. *Med Phys.* 2022;49(1):343-356.
- [12] Liu X, Xing F, Prince JL, et al. Dual-cycle constrained bijective VAE-GAN for TAGGED-TO-CINE magnetic resonance image synthesis [J]. *Proc IEEE Int Symp Biomed Imaging.* 2021; 10:1109
- [13] Yan S, Wang C, Chen W, et al. Swin transformer-based GAN for multi-modal medical image translation[J]. *Front Oncol.* 2022; 12:942511.
- [14] Zhou T, Fu H, Chen G, et al. Hi-Net: Hybrid-Fusion Network for Multi-Modal MR Image Synthesis[J]. *IEEE Trans Med Imaging.* 2020 ;39(9):2772-2781.
- [15] 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 ;24;7(2).
- [16] Klages P, Benslimane I, Riyahi S, et al . Patch-based generative adversarial neural network models for head and neck MR-only planning[J]. *Med Phys.* 2020 ;47(2):626-642.
- [17] Hsu SH, Han Z, Leeman JE, et al. Synthetic CT generation for MRI-guided adaptive radiotherapy in prostate cancer[J]. *Front Oncol.* 2022;23;12:969463.
- [18] Shin, H.C., Ihsani,et al. Generative adversarial networks with bidirectional encoder representations from transformers for mri to pet synthesis.2020;2008:04393.
- [19] Lin H, Figini M, D'Arco F, et al. Low-field magnetic resonance image enhancement via stochastic image quality transfer[J]. *Med Image Anal.* 2023 ;87:102807.
- [20] Liu J, Pasumarthi S, Duffy B, et al. One Model to Synthesize Them All: Multi-Contrast Multi-Scale Transformer for Missing Data Imputation[J]. *IEEE Trans Med Imaging.* 2023;42(9):2577-2591.
- [21] Schwab J, Antholzer S, Haltmeier M. Big in Japan: Regularizing Networks for Solving Inverse Problems[J]. *J Math Imaging Vis.* 2020;62(3):445-455.
- [22] Häggström I, Schmidlein CR, Campanella G, et al. DeepPET: A deep encoder-decoder network for directly solving the PET image reconstruction inverse problem[J]. *Med Image Anal.* 2019 ;54:253-262.
- [23] Wang X, Yang M, Tosun S, et al. CalDiff: calibrating uncertainty and accessing reliability of diffusion models for trustworthy lesion segmentation[J]. *IEEE J Biomed Health Inform.* 2025;23;PP.
- [24] Eshraghi Boroojeni P, Chen Y, Commean PK, et al. Deep-learning synthesized pseudo-CT for MR high-resolution pediatric cranial bone imaging (MR-HiPCB)[J]. *Magn Reson Med.* 2022;88(5):2285-2297.
- [25] Qi M, Li Y, Wu A, 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.

Medical Image Synthesis Based on Deep Learning: Methods, Applications and Future Directions

Yang Li

Chongqing Medical University Affiliated Rehabilitation Hospital, Chongqing, China

Abstract: With the rapid advancement of medical imaging technology, deep learning has demonstrated tremendous potential in medical image synthesis, offering new opportunities for clinical diagnosis, treatment planning, and medical research. This paper systematically reviews deep learning-based medical image synthesis methods, focusing on the application of Generative Adversarial Networks (GANs) in cross-modal synthesis and domain adaptation, the high-fidelity generation capabilities of Diffusion Models (DM) and their limitations in multimodal synthesis (such as insufficient diversity of synthesized samples), as well as the advantages of Transformer models, hybrid models, and domain adaptation techniques. In terms of application scenarios, the paper analyzes practical cases including cross-modal image synthesis (e.g., MRI to CT, MRI to PET conversion), intra-modal image enhancement and standardization, image reconstruction and denoising, longitudinal image generation and disease progression modeling, and specific lesion or tissue synthesis. It also highlights the limitations of existing methods in generalization, computational efficiency, and ethical compliance. Future research should focus on improving model robustness, multimodal fusion capabilities, and clinical applicability. This study provides a comprehensive reference for innovation and clinical translation in the field of medical image synthesis.

Keywords: Deep learning; Medical image synthesis; Generative adversarial networks (GANs); Diffusion models (DMs); Cross-modal synthesis