

生成模型在胸部X射线图像中的应用综述

储欢, 肖琪, 王慧玲*

伊犁师范大学网络安全与信息技术学院, 新疆 伊宁

收稿日期: 2025年3月16日; 录用日期: 2025年4月16日; 发布日期: 2025年4月24日

摘要

生成模型在医学影像领域的快速发展为胸部X射线(CXR)图像的合成、编辑与增强等领域提供了新的技术手段。本文系统综述了生成对抗网络(GAN)、变分自编码器(VAE)及扩散模型在CXR图像中的研究进展与应用。VAE通过隐变量学习生成数据分布, 在疾病检测与图像重构中表现稳健, 但生成图像常存在模糊现象; GAN凭借其高真实感图像生成能力, 被广泛用于解决数据稀缺问题及跨模态图像合成, 但其训练不稳定性及模式崩溃问题仍需优化; 扩散模型凭借逐步去噪的生成机制, 在图像质量与多样性上展现出超越GAN的潜力, 成为当前研究热点。文章进一步分析了该领域的研究现状, 总结了生成模型在数据增强、图像生成、及图像编辑中的创新应用, 并对比了不同技术的优势与局限性。尽管生成模型在提升诊断效率与数据隐私保护方面成果显著, 但仍面临伦理、法律及模型泛化性等挑战。未来研究需聚焦多模态生成、隐私保护框架设计及病理特征解耦, 以推动生成模型在临床中的实际应用。本文为医学影像领域的研究者提供了技术参考与方向指引, 具有重要的学术价值与应用前景。

关键词

VAE, GAN, 扩散模型, 胸部X射线

A Review of the Application of Generative Models in Chest X-Ray Images

Huan Chu, Qi Xiao, Huiling Wang*

School of Network Security and Information Technology, Yili Normal University, Yining Xinjiang

Received: Mar. 16th, 2025; accepted: Apr. 16th, 2025; published: Apr. 24th, 2025

Abstract

The rapid development of generative models in the field of medical imaging has provided new

*通讯作者。

文章引用: 储欢, 肖琪, 王慧玲. 生成模型在胸部 X 射线图像中的应用综述[J]. 计算机科学与应用, 2025, 15(4): 287-300.
DOI: 10.12677/csa.2025.154101

technical means for the synthesis, editing and enhancement of chest X-ray (CXR) images. This paper systematically reviews the research progress and applications of generative adversarial networks (GANs), variational autoencoders (VAEs) and diffusion models in CXR images. VAEs generate data distributions through latent variable learning, and are robust in disease detection and image reconstruction, but the generated images are often blurred; GANs are widely used to solve data scarcity problems and cross-modal image synthesis due to their ability to generate highly realistic images, but their training instability and mode collapse problems still need to be optimized; diffusion models have the potential to surpass GANs in image quality and diversity due to their step-by-step denoising generation mechanism, and have become a current research hotspot. This paper further analyzes the current research status in this field, summarizes the innovative applications of generative models in data enhancement, image generation, and image editing, and compares the advantages and limitations of different technologies. Although generative models have achieved remarkable results in improving diagnostic efficiency and protecting data privacy, they still face challenges such as ethics, law, and model generalization. Future research needs to focus on multimodal generation, privacy protection framework design, and pathological feature decoupling to promote the practical application of generative models in clinical practice. This article provides technical references and direction guidance for researchers in the field of medical imaging, and has important academic value and application prospects.

Keywords

VAE, GAN, Diffusion Model, Chest X-Ray

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

自 1895 年 Rontgen 发现 X 射线以来, X 射线成像技术便迅速成为医学诊断的重要工具, 尤其在对人体疾病的检测及诊断中扮演着愈发重要的角色。胸部 X 射线图像能够清晰地显示肺部、心脏、肋骨以及胸膜等相关解剖结构图像, 对于肺炎、肺结核、肺癌、气胸等多种胸部疾病的筛查过程中有着不可或缺的作用。随着图像生成技术的不断进步, 生成模型在图像生成和处理领域的应用日益广泛, 为胸部 X 射线图像的分析与诊断提供了新的视角和方法[1]。

近年来, 生成模型在深度学习领域取得了显著的进展, 为数据生成和模拟提供了强大的工具。最初, Kingma 和 Welling 于 2013 年提出了变分自编码器(Variational Autoencoder, VAE) [2], VAE 是一种基于概率的生成模型, 其通过引入潜在变量来捕捉数据的高维特征, 从而生成与原始数据相似的新样本。在胸部 X 射线图像的应用中, VAE 能够学习到图像的潜在分布, 进而生成具有多样性和逼真度的胸部 X 射线图像。这种生成能力对于数据增强、病变模拟以及医学教育等方面具有重要意义[3]。例如, 通过 VAE 生成的胸部 X 射线图像, 医生可以更加直观地了解不同病变类型下的图像特征, 从而提高诊断的准确性。随后, Goodfellow 等人于 2014 年提出了生成对抗网络(Generative Adversarial Networks, GAN) [4], GAN 则由生成器和判别器两个网络组成, 通过相互对抗训练来不断提升生成图像的质量。在胸部 X 射线图像的应用中, GAN 同样展现出出色的图像生成能力。与 VAE 不同, GAN 更侧重于生成图像与真实图像之间的细微差别, 使得生成的图像在细节上更加逼真[5]。这种精细的图像生成能力对于医学图像的分割、病变检测以及诊断辅助等方面具有显著优势。通过 GAN 生成的胸部 X 射线图像, 医生可以更加准确地定位病变位置, 提高诊断的效率和准确性[6]。

然而, VAE 和 GAN 在生成高分辨率和细节丰富的图像时仍存在一定的局限性。扩散模型的出现一定程度上解决了这些问题。扩散模型(Diffusion Models) [7]最早由 Sohl-Dickstein 等人于 2015 年提出。随后, Ho 等人进一步改进了该模型, 提出了去噪扩散概率模型(Denoising Diffusion Probabilistic Models, DDPM) [8]。扩散模型则是一种基于概率扩散过程的生成模型, 其通过逐步添加和去除噪声, 学习数据分布, 从而生成高质量和高多样性的图像。随着扩散模型的发展, 其已经在医学领域高质量图像合成方面显示出极具前景的生成能力[9] [10]。在生成高分辨率和细节丰富的图像方面, Rombach 等人提出了一种名为潜在扩散模型(Latent Diffusion Models, LDMs) [9]的方法, 通过在预训练自编码器的潜在空间中应用扩散模型, 显著提高了生成高分辨率图像的效率和质量。实验结果表明, LDMs 在图像生成任务中(如图像修复、语义场景合成和超分辨率)表现出色, 尤其在高分辨率和细节丰富度方面优于传统方法。在胸部 X 射线图像的应用中, 扩散模型能够捕捉到图像中的细节信息和纹理特征, 从而生成具有高度逼真度和多样性的图像。这种生成能力对于医学图像的增强、去噪以及超分辨率重建等方面具有重要意义[10]。通过扩散模型处理的胸部 X 射线图像, 其质量和清晰度能够得到显著提升, 有助于医生更加准确地观察和分析图像中的病变特征。例如 Anand 等人提出了一种基于深度对比扩散学习的网络(Deep Contrast Diffusion Network, DCDN) [11], 专门用于增强胸部 X 射线(CXR)图像的对比度和清晰度, 实验结果表明, DCDN 能够显著提升 CXR 图像的质量, 尤其是在低对比度区域, 使肺部病变特征(如结节、浸润等)更加明显。扩散模型的最新进展还促成了如 RoentGen [12]等复杂工具的开发, 这是一种视觉-语言基础模型, 能够根据放射学报告生成高质量 CXR 图像。其详细描述了如何通过适应预训练的潜在扩散模型, 克服自然图像与医疗图像的分布差异, 生成多样且视觉上逼真的 CXR 图像。这种能力不仅支持数据增强, 还通过文本相关性增强了图像特征的可解释性, 特别适用于研究和教育目的。

此外, 近期研究如 ViewXGen [13]探索了生成特定视图的 CXR 图像, 标志着领域的重要进展。这些模型通过设计特定视图的标记, 满足用户偏好, 进一步扩展了生成模型的应用场景。

尽管生成模型展现出巨大潜力, 其在临床实践中的采用仍面临多重挑战。一大难题是确保生成图像的现实性和准确性, 任何偏差可能导致误诊或误导训练模型。此外, 模型可解释性是一个关键问题, 特别是在医疗领域, 医生需要理解模型如何生成特定输出以建立信任。其次, 伦理考虑也至关重要, 例如数据隐私和合成图像的潜在误用。生成模型是在真实的患者数据上训练的, 可能被用于创建侵犯隐私的图像, 或用于恶意目的。最后, 准确的多标签病理标注图像对于肺部疾病的诊断和深度学习模型的训练极其重要。然而, 在实际应用中, 往往难以获取足够数量和高质量的胸部 X 射线图像数据集, 这限制了生成模型性能的进一步提升。同时由于患者隐私、数据获取和标注成本等原因, 更加提升了获得大量高质量的多病理标签 X 射线图像的难度。因此, 未来研究应聚焦于开发生成图像的鲁棒评估指标, 通过可解释 AI 技术增强模型透明度, 并实施严格的数据治理政策。

总之, 生成模型已成为 CXR 图像分析中不可或缺的工具, 提供了解决数据稀缺和高质量训练数据需求的长久挑战的解决方案。随着这些模型的持续发展, 其应用预计将扩展, 有望改变胸部疾病的检测和诊断方式。本文后续部分将全面回顾生成模型在 CXR 图像应用的文献, 涵盖其历史发展、当前应用和未来方向。

2. 生成模型简介

生成模型是一类重要的机器学习模型, 其目标是通过学习数据的潜在分布, 生成与训练数据相似的新样本。近年来, 生成模型在图像处理、自然语言处理和医学成像等领域展现出了卓越的性能。特别是在医学图像分析中, 生成模型为应对数据稀缺、隐私保护和图像质量提升等挑战提供了有效的解决方案 [14]。本文将详细介绍三种主流生成模型的基本原理和特点: 变分自编码器(VAE)、生成对抗网络(GAN)和扩散模型。

2.1. 变分自编码器(VAE)

变分自编码器(Variational Autoencoder, VAE)是一种基于变分推断的生成模型, 由 Kingma 和 Welling 在 2013 年首次提出[2]。其核心目标是通过输入数据 x 的潜在变量 z 进行变分推断, 学习数据的低维隐空间(Latent Space)表示, 然后通过解码器从潜在空间中采样, 生成新的数据样本。其训练目标是最大化数据的对数似然, 同时利用变分推断来近似潜在变量的后验分布。

VAE 的结构主要包括以下两部分:

编码器(Encoder): 将输入数据 $x \in \mathbb{R}^D$ 映射至隐变量 $z \in \mathbb{R}^d$ ($d \ll D$) 的变分后验分布。假设隐空间服从各向同性多元高斯分布, 编码器输出参数为:

$$q_{\phi}(z|x) = \mathcal{N}(z; \mu_{\phi}(x), \text{diag}(\sigma_{\phi}^2(x)))$$

其中 ϕ 为编码器参数, μ_{ϕ} 和 σ_{ϕ}^2 分别表示均值向量和对角协方差矩阵。

解码器(Decoder): 解码器定义生成过程 $p_{\theta}(x|z)$, 通过隐变量 z 重构数据分布。对于图像数据, 通常建模为伯努利分布(二值像素)或高斯分布(连续像素):

$$p_{\theta}(x|z) = \prod_{i=1}^D \mathcal{N}(x_i; \mu_{\theta,i}(z), \sigma_{\theta,i}^2(z))$$

其中 θ 为解码器参数。

VAE 的训练通过优化一个损失函数实现, 该损失函数由两部分组成:

重构损失(Reconstruction Loss): 衡量生成样本与输入数据的相似度, 鼓励模型生成接近原始数据的输出。

KL 散度(Kullback-Leibler Divergence): 正则化隐空间分布 $q_{\phi}(z|x)$ 使其逼近先验分布 $p(z) = \mathcal{N}(\mathbf{0}, I)$, 其闭合解为:

$$D_{\text{KL}} = \frac{1}{2} \sum_{i=1}^d (\sigma_{\phi,i}^2 + \mu_{\phi,i}^2 - 1 - \log \sigma_{\phi,i}^2)$$

损失函数的数学表达式为:

$$\mathcal{L}(\theta, \phi; x) = \mathbb{E}_{q_{\phi}(z|x)} [\log p_{\theta}(x|z)] - D_{\text{KL}}(q_{\phi}(z|x) \| p(z))$$

VAE 在图像重建、去噪和数据增强等任务中表现出色, 尤其适用于需要学习数据潜在表示的场景[15]。

2.2. 生成对抗网络(GAN)

生成对抗网络(Generative Adversarial Networks, GAN)由 Goodfellow 等人在 2014 年首次提出, 是一种基于博弈论框架的生成模型, 通过对抗训练机制实现数据分布建模[4]。GAN 包含两个核心组件:

生成器(Generator): 定义从潜在空间(latent space)到数据空间的映射函数 $G: \mathcal{Z} \rightarrow \mathcal{X}$, 其输入为服从先验分布 $p_z(z)$ (通常为标准正态分布或均匀分布)的随机噪声向量 $z \in \mathcal{Z}$, 输出为合成数据样本 $G(z)$ 。生成器的优化目标是使合成数据分布 p_g 逼近真实数据分布 p_{data} 。

判别器(Discriminator): 定义判别函数 $D: \mathcal{X} \rightarrow [0, 1]$, 其输入为真实数据样本 $x \sim p_{\text{data}}$ 或生成样本 $G(z)$, 输出为输入样本属于真实分布的概率估计。判别器的目标是构建数据分布差异的度量标准。GAN 的训练过程是一个博弈过程, 生成器和判别器通过以下对抗性损失函数相互竞争:

$$\min_G \max_D \mathbb{E}_{x \sim p_{\text{data}}} [\log D(x)] + \mathbb{E}_{z \sim p_z} [\log (1 - D(G(z)))]$$

在这个过程中, 判别器试图最大化损失函数以正确区分真假数据, 生成器通过最小化判别器的鉴别准确率驱动合成分布 p_g 向 p_{data} 收敛以欺骗判别器。经过充分训练, 生成器能够逐渐逼近真实数据的分布, 生成高质量的样本。GAN 在图像生成、风格迁移和数据增强等领域有着广泛应用, 尤其在生成逼真图像方面表现尤为突出[16]。

2.3. 扩散模型

扩散模型(Diffusion Models)是一种近年来迅速发展的生成模型, 其理论体系建立在非平衡态热力学(Non-equilibrium Thermodynamics)与随机微分方程(Stochastic Differential Equations)的基础之上。该模型通过构建参数化的马尔可夫链(Markov Chain), 实现了对数据分布的高效建模。Ho 等人在 2020 年提出的去噪扩散概率模型(Denoising Diffusion Probabilistic Models, DDPM)是扩散模型的代表性工作, 其通过引入渐进式加噪 - 去噪机制, 为扩散模型的实际应用奠定了理论基础[8]。

扩散模型的工作原理分为两个阶段:

前向扩散过程: 给定原始数据样本 $x_0 \sim q(x_0)$, 前向过程定义为具有固定参数的马尔可夫链:

$$q(x_{1:T} | x_0) := \prod_{t=1}^T q(x_t | x_{t-1})$$

其中每个转移核服从高斯分布:

$$q(x_t | x_{t-1}) := \mathcal{N}(x_t; \sqrt{1 - \beta_t} x_{t-1}, \beta_t \mathbf{I})$$

式中 $\beta_t \in (0, 1)$ 为预设的噪声调度系数。当 $T \rightarrow \infty$ 时, x_T 依分布收敛于各向同性高斯分布 $\mathcal{N}(\mathbf{0}, \mathbf{I})$ 。

逆向去噪过程: 逆过程通过可学习参数 θ 构建变分分布:

$$p_\theta(x_{0:T}) := p(x_T) \prod_{t=1}^T p_\theta(x_{t-1} | x_t)$$

其中逆转转移核参数化为:

$$p_\theta(x_{t-1} | x_t) := \mathcal{N}(x_{t-1}; \mu_\theta(x_t, t), \Sigma_\theta(x_t, t))$$

通过优化证据下界(Evidence Lower Bound, ELBO):

$$\mathbb{E}_q[\log p_\theta(x_0)] \geq \mathbb{E}_q\left[\log \frac{p_\theta(x_{0:T})}{q(x_{1:T} | x_0)}\right]$$

推导可得训练目标函数:

$$\mathcal{L}_{\text{ELBO}} = \mathbb{E}_{t, x_0, \epsilon} \left[\|\epsilon - \epsilon_\theta(x_t, t)\|_2^2 \right]$$

式中 ϵ_θ 为噪声预测网络, t 均匀采样于 $\{1, \dots, T\}$ 。

相比于 VAE 和 GAN, 扩散模型在生成高分辨率、细节丰富的图像方面具有显著优势, 尤其在医学图像生成等任务中展现出巨大潜力[17]。

3. 基于变分自编码器(VAE)的胸部 X 光片生成与编辑

3.1. VAE 的理论基础与早期医学影像应用

变分自编码器(VAE)的提出为医学影像生成与编辑奠定了基础。2013 年, Kingma 与 Welling 首次提出 VAE 框架, 通过引入概率潜在空间和变分推断, 实现了数据分布的隐式建模[2]。早期研究主要聚焦于 VAE 在医学影像重建任务中的基础能力。例如, Park 等人将变分自编码器(VAE)应用于医学图像分析中

的流形学习,通过对 OASIS 数据集的实验,展示了 VAE 在学习低维流形以用于医学图像生成和临床参数分类方面的潜力[18]。然而,标准 VAE 在胸部 X 光片生成中存在明显局限性:生成图像模糊、细节丢失严重,且难以捕捉病理特征的复杂分布[19]。

为提升生成质量,研究者开始探索改进架构。2016 年, Higgins 等人提出 β -VAE,通过调整 KL 散度项的权重 β ,增强潜在空间的可解释性[20]。这一改进鼓励模型在潜在空间中进行解纠缠因子学习,能发现可解释的因子化表示[21]。同时,对抗自编码器(AAE)的引入为生成任务提供了新思路。Makhzani 等人(2016)通过对抗训练约束潜在空间分布,在胸部 X 光片生成中实现了更清晰的纹理细节[19]。

3.2. VAE 在肺部分割与重建中的突破

随着医学影像数据量的增加,VAE 被广泛用于肺部区域的精确分割与重建。Selvan 等人(2021)提出基于 VAE 的编码器-解码器架构,针对高不透明度胸部 X 光片,实现了 88.15%的分割准确率,显著提升了复杂病例的处理能力[22]。同年, Crespi 等人对比了 DenseNet121、DenseNet201 等模型结合 VAE 的重建效果,发现不同模型在保留肺部解剖结构上存在差异,为生成质量的优化提供了方向[23]。

在分割算法优化方面,2021 年 Cao 等人提出的“融合 VAE 与三端注意力机制”模型[24]通过引入通道和空间注意力模块,将肺部分割的准确率提升至 97.8%,证明了 VAE 与注意力机制结合的优越性。此外, Gerlings 等人[25]展示了 U-Net 与 VAE 联合分割肺区的完整流程,包括肺掩膜生成和异常区域评估,成为 COVID-19 肺炎检测的重要工具。

3.3. VAE 生成与编辑技术的精细化发展

在病理生成方面, Bercea 等人(2022)提出结合全局图像先验与局部形变场的可变形自编码器 MorphAEus,并系统评估了与 VAE、 β -VAE 和 S-Intro VAE 等模型生成胸部 X 光片的效果对比,揭示传统自编码器在无监督异常检测中的局限性,所提方法在控制生成图像的病理特征上更具灵活性[19]。此类研究为合成数据增强提供了技术支持,尤其在 COVID-19 数据稀缺的场景下(如 Chatterjee 等人中利用 VAE 平衡数据集,提升分类器性能[26])。

在胸部 X 射线编辑方面,核心挑战在于实现局部病理特征的精准控制。Montenegro 等人(2023)提出解耦 VAE 架构,将胸部 X 光片分解为医学特征、身份特征和其他特征三个独立潜在空间,实现了患者身份信息与病理特征的分离编辑[27]。在分割-编辑联合任务中, Chambon 等人(2022)开发融合三端注意力机制的 VAE-U-Net 模型,肺野分割的 Jaccard 指数达到 0.92,同时支持基于块掩码的肺水肿模拟[24]。

在反事实生成方面, Gu 等人(2023)进一步将 VAE 与扩散模型结合,开发的 BiomedJourney 框架支持基于文本指令的反事实生成,如在胸膜积液(Pleural Effusion)的病理分类 AUC 指标上, BiomedJourney 达到 88.11(对比 RoentGen 的 86.20),表明其在该类指令上的准确性显著优于基线[28]。

3.4. VAE 混合模型创新

近年研究趋向于将 VAE 与其他深度学习模型结合。VAE 和生成对抗网络(GAN)是两种在图像生成领域中广受关注的模型。VAE 擅长捕捉数据的潜在分布并生成多样化的样本,但其生成的图像往往模糊且缺乏细节。而 GAN 则以其生成逼真图像的能力著称,但训练过程可能不稳定,且容易出现模式崩塌问题(mode collapse)。因此,将 VAE 与 GAN 结合成混合模型可以发挥二者的优势,实现既多样又逼真的图像生成。例如, Li 等人(2021)提出了一种新的 VAE-GAN 模型,用于从结构 MRI 中合成动脉自旋标记(ASL)图像[29]。该模型将 VAE 作为生成器,并结合 GAN 的对抗训练机制,成功生成高质量的 ASL 图像。尽管该研究针对的是 MRI,但其方法论可以推广到其他医学影像领域,包括胸部 X 光片。该模型的

关键创新在于通过 VAE 捕捉数据的潜在分布, 并通过 GAN 的判别器确保生成图像的真实性。

除了与 GAN 结合, 在图像生成领域, VAE 也常与去噪扩散网络结合以增强生成图像的清晰度与病理特征的可控性。例如, Chambon 等人提出的 RoentGen 模型(2022) [12], 通过“条件去噪 U-Net + VAE 解码器”架构实现了从文本到胸部 X 光图像的端到端生成。具体地, 模型首先通过随机高斯噪声模块引入数据扰动, 随后利用条件去噪 U-Net 对噪声进行分层过滤, 最终通过 VAE 解码器重构高分辨率图像。该模型支持输入如“双侧胸腔积液和轻度肺水肿”等复杂医学描述, 生成的图像在肺部纹理和病变区域分布上接近真实影像, 验证了多模态输入的可行性。进一步地, VAE 与扩散模型的融合便成为了新的趋势。Siddiqui 等人(2024)开发的 VALD-MD 框架[30], 将 VAE 编码器与潜扩散模型结合, 通过医学文本提示指导生成健康或病理对比图像, 创新性地实现解剖结构一致性保留与零样本疾病诱导, 实验表明其在图像质量和诊断解释性上显著优于传统 GAN 方法。

4. 基于生成对抗网络(GAN)的胸部 X 光片生成与编辑

4.1. GAN 的理论基础与早期医学影像应用

生成对抗网络(Generative Adversarial Networks, GANs)由 Goodfellow 等人于 2014 年首次提出, 其核心思想是通过生成器(Generator)与判别器(Discriminator)的对抗训练实现数据分布建模[31]。在医学图像领域, GAN 因其强大的图像生成能力被广泛应用于数据增强、图像修复和病理模拟。例如, 刘建伟等(2019)系统总结了 GAN 的变体, 包括条件生成对抗网络(cGAN)、Wasserstein GAN (WGAN)和循环一致生成对抗网络(CycleGAN), 这些变体通过改进损失函数或网络结构, 显著提升了医学图像的生成质量[32]。Abeer Aljohani (2022)进一步指出, GAN 在生成肺部 CT、脑部 MRI 和皮肤病变图像中表现出色, 其生成的肺结节样本在形态学特征上与真实数据高度一致[33]。

4.2. GAN 的变体在胸部 X 光片生成与编辑中的发展

随着 GAN 的技术迭代, 越来越多的胸部 X 光片生成研究开始转向基于 GAN 的变体架构。例如, Venu 等人(2020)采用深度卷积生成对抗网络(DCGAN)生成胸部 X 光片, 但存在不稳定性和难以达到理论最优解的问题, 可能导致生成质量受限[34]。为提升生成效果, 后续研究引入条件生成机制。证据显示, Dhawan 等人(2024)提出了一种基于 AC-GAN (Auxiliary Classifier GAN)的条件生成框架, 通过输入类别标签(如健康、病毒性肺炎、细菌性肺炎)生成特定病理状态的合成医学影像(胸部 X 光和脑 MRI), 并验证了合成数据在提升分类模型(如 EfficientNet v2)性能中的有效性[35]。近年来, Pix2Pix 和 CycleGAN 成为主流框架。Huang 等人(2022)使用 Pix2pixGAN 结合互补双视角(0° & 180°)在胸部 X 光数据上实现了高精度的透视变形校正, SSIM 达到 0.89 以上, 显著优于单视角输入(类似传统单视图方法) [36]。然而, Pix2Pix 需严格配对的训练数据, 而 CycleGAN 则突破此限制。例如, Liang 等人(2022)提出了一种基于自适应循环生成对抗网络(Ad CycleGAN)的新方法, 通过在传统 CycleGAN 架构中引入独立预训练准则, 实现了正常胸部 X 光图像与 COVID-19 胸部 X 光图像的转换, 实验表明该方法显著提升了生成图像的 FID 分数和分类准确率, 有效改善了医学图像生成质量与类别不平衡问题[37]。同时, Kong 等人(2021)提出了一种名为 RegGAN 的新型无监督医学图像转换模式, 通过引入配准网络自适应校正未对齐数据的噪声分布, 在无需严格像素对齐的情况下, 显著优于现有监督 Pix2Pix 和无监督 CycleGAN 方法, 实现了跨模态医学图像转换的高精度与强鲁棒性[38]。

4.3. GAN 在胸部 X 光片生成与编辑中的局限性

在胸部 X 射线病变区域的添加与去除领域, 大多倾向于异常到正常转换[6] [37] [39] [40]、单标签编

辑[41]-[43]、少量标签编辑[44]-[46]。异常到正常转换的研究中, Tang 等人(2021)提出了一种 GAN 的变体, 通过合成正常图像来对胸部 X 射线图像中的疾病进行分解。这一方法通过生成无病变的图像, 使病变特征更加明显[6]。Qin 等人受 Tang 等人[6]的启发, 利用 GAN 的变体将输入图像转换为正常图像, 提出了疾病分解网络(DDN)架构和组成组件, 以提高翻译图像的质量和疾病定位性能[39]。Kim 等人(2023)通过在具有相似外观的胸部 X 射线图像中寻找正常局部区域并应用融合过程来消除异常区域。然后, 利用这些合成的匹配对图像训练 GAN, 以构建将给定胸部 X 射线转化为其对应正常胸部 X 射线的翻译模型[40]。以上模型仅关注将异常胸部 X 射线转换成正常, 不能对胸部 X 射线添加病理。单标签编辑的研究中, Jin 等人(2023)提出了使用条件循环生成对抗网络(Conditional Cycle GANs)生成肺炎进展的胸部 X 光片图像的方法[42]。该方法通过在正常和肺炎图像之间进行域转换, 生成中间域的图像, 从而模拟肺炎的进展过程。Saboo 等人(2021)在 CheXpert 数据集[47]上训练 StyleGAN 合成 X 射线图像, 通过引导潜在代码的操作, 将病理特征赋予健康患者的 X 光片, 但仅用于心脏轮廓操作[41]。Liang 等人(2020)提出了通过图像到图像的 GAN 翻译来增强自动化 COVID-19 胸部 X 光片诊断的方法[43], 该方法利用 GAN 将输入的胸部 X 光片转换为 COVID-19 胸部 X 光的图像。这些研究都只关注于肺部的一种病症, 存在一定的局限性。在少量标签编辑的研究中, Saboo 等人(2021)通过在预训练的 StyleGAN 的潜在流形上发现最佳轨迹, 能够在用户定义的感兴趣区域内生成疾病进展的图像, 但其仅关注于心脏增大, 肺浑浊病理的编辑[45]。Wang 等人(2022)提出了基于肺部分割重建的数据增强方法[46], 能够生成基于异常图像的健康参考图像, 其仅关注于实变和胸腔积液病理的编辑。Weber 等人(2022)提出了一种通过 GAN 反转生成高分辨率胸部 X 射线隐式嵌入的方法[44], 利用 GAN 模型的反转能力, 将原始图像嵌入到潜在空间中, 然后通过调整潜在向量来生成高质量的合成图像, 但仅能实现对全图像有很大影响的病理的编辑, 如肺不张、胸腔积液或心脏肿大。以上的研究中, 缺乏对胸部 X 射线图像进行多种病理标签编辑的研究, 因为生成对抗网络(GANs)针对不同疾病语义的受控生成是困难的, 其难以将局部不同疾病语义因素从全局疾病无关因素中分离出来。

5. 基于扩散模型(Diffusion Model)的胸部 X 光片生成与编辑

5.1. 扩散模型基本原理及其医学图像生成优势

扩散模型(Diffusion Model)作为一种基于马尔可夫链的概率生成模型, 其核心思想是通过逐步添加噪声破坏原始数据分布, 再通过反向过程学习数据分布重建规律。在医学图像生成领域, 扩散模型展现出显著优势: 相较于生成对抗网络(GANs), 其训练稳定性更高且模式崩溃风险更低。例如, Khader 等人(2023)在 3D 医学图像生成任务中证实, 扩散模型在仅使用 1000 个样本的小型数据集时仍能实现稳健收敛, 其生成的脑磁共振图像 MS-SSIM 分数(0.8557)更接近真实数据分布(0.8095), 而 WGAN-GP 生成的图像因模式崩溃导致 MS-SSIM 高达 0.9996 [48]。这一特性使得扩散模型在保留病理细节的同时, 能生成具有临床意义的图像多样性, 为合成数据增强提供了新范式。

5.2. 早期胸部 X 光片生成研究

在扩散模型理论框架尚未成熟的阶段, 研究者已尝试将其应用于胸部 X 光片生成。Ho 等人(2020)提出的去噪扩散概率模型(DDPM)首次在自然图像生成中实现质量突破, 其在无条件 CIFAR10 数据集上取得了 3.17 的 FID 分数, 超越同期 GANs 模型, 为医学图像生成奠定算法基础[8]。早期医学应用集中于数据增强场景, 例如 Schaudt 等人(2023)在肺炎 X 光片分类任务中, 使用扩散模型生成合成数据以平衡类别分布, 显著提升了数据稀少类别的 F1-score 分数[49]。然而, 此时生成的图像分辨率较低, 且缺乏明确的

病理控制机制。

5.3. 技术改进阶段

此阶段研究重点转向提高生成图像的分辨率与临床可解释性。Packhäuser 团队(2022)开发了基于潜在扩散模型(Latent Diffusion Model)的匿名化框架,并通过隐私增强的采样策略来防止生物特征泄露。成功生成 256×256 分辨率的匿名胸部 X 光片,其生成的合成数据在胸腔异常分类任务中达到 ROC-AUC 0.89,与真实数据性能差距仅 3.5% [50]。Gu 等人(2023)提出的 BiomedJourney 模型,通过融合 BiomedCLIP 文本编码器与潜在扩散模型(LDM),首次实现了基于自然语言指令的高质量反事实医学图像生成,在 MIMIC-CXR 数据集上病理准确性(80.54 AUC)和患者特征保留(97.22 种族分类 AUC)均显著超越现有方法[28]。

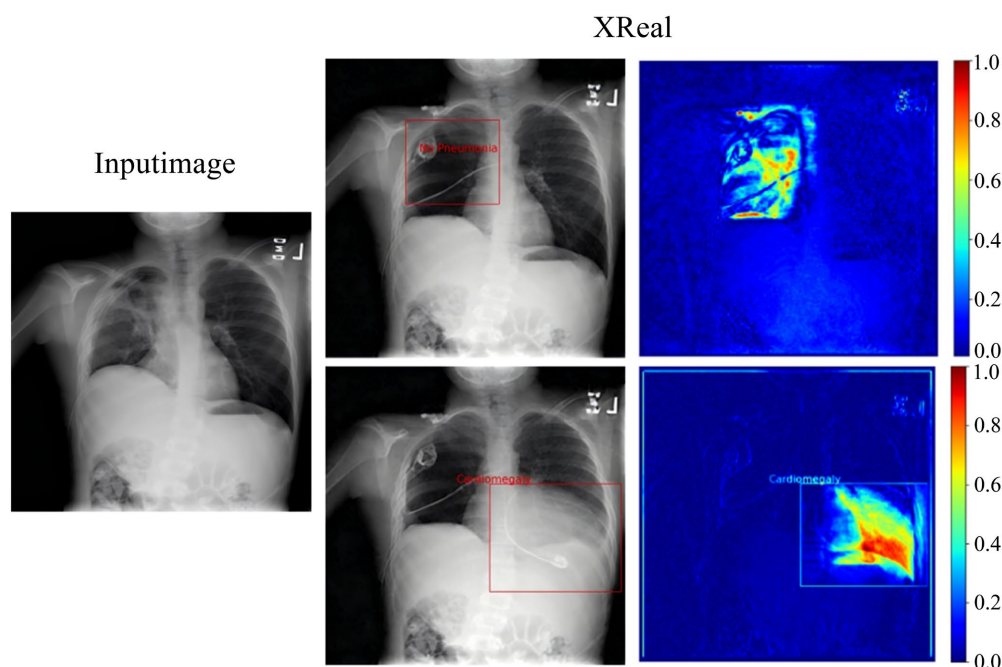


Figure 1. The image provided by XReal and the residual with the original image
图 1. XReal 中提供的图像以及与原图像的残差

5.4. 生成与编辑中的发展

在胸部 X 光片编辑领域,早期研究主要依赖 GAN 生成胸部 X 光片。然而 GAN 存在模式崩塌问题,难以生成精细解剖结构。Dhariwal 等人的研究表明,与 GANs 相比扩散模型可以实现优于当前最先进的生成模型的图像样本质量[51],于是越来越多的领域内学者们开始关注于使用扩散模型对胸部 X 射线进行编辑[52]-[55]。其中 Wolleb 等人使用 Dhariwal 等人[51]中提出的带有分类器引导的方法并使用 DDIM [56]中提出的确定性采样方案,将病理的胸部 X 射线图像转换为健康的胸部 X 射线图像,通过比较原始图像和生成的正常图像之间的差异来检测病变区域,且此研究仅关注于胸膜积液[53]。Fathi 等人通过在推理过程中利用外部预训练的二元伪影检测器,指导基于扩散的反事实图像生成器,以实现更准确的解释性,但其编辑图像仅产生了微小的变化且仅关注于胸腔积液的编辑[54]。Weber 等人提出了一种级联 LDM Cheff [52],可以生成 100 万像素尺度的高质量胸部 X 射线图像,并允许通过文本提示对扩散过程进行细粒度控制,从而实现多种病理的胸部 X 射线的生成。Hashmi 等人在 Weber 等人[52]的基础上通过

对生成图像的解剖和病理施加空间控制实现了胸部器官和病理的定向编辑[55]，其虽然实现了多标签的病理编辑但忽视了编辑区域胸部解剖结构的一致性。如图 1 所示，最左边的图像(输入图像)是显示左肺有肺炎的胸部 X 光片。最右边的图像是 XReal 方法生成的结果以及残差图像，突出显示了输入和 XReal 生成的图像之间的差异。在病理去除方面，XReal 方法对肺部结构进行了重大修改，无法保留原始骨骼结构，并在生成的图像中添加了新的支撑装置。对于病理添加，XReal 方法同样大幅改变肺部结构，无法保留原始骨骼结构，引入了新的支撑装置，导致肺叶尖端消失。此外，XReal 添加的心脏肥大程度显得夸张。这使其不适用于医学临床研究，限制了其在放射科医生培训和反事实诊断中的应用。

6. 生成模型优缺点分析

基于本论文各章节对生成模型技术特点的系统分析，并结合当前研究进展与实验结果，本文对主要生成模型在胸部 X 射线图像应用中的优缺点进行了对比总结，如表 1 所示：

Table 1. Comparison of different model methods
表 1. 不同模型方法比较

模型	优势	局限性	典型应用场景
VAE	1) 潜在空间可解释性强 2) 支持半监督学习框架集成 3) 训练过程稳定	1) 生成图像清晰度较低 2) 对复杂病理表征能力有限 3) 存在后验坍塌风险	疾病解耦分析 低剂量图像重建
GAN	1) 生成图像细节丰富，可模拟真实病理特征(如气胸、COVID-19病变) 2) 支持无配对数据增强(CycleGAN) 3) 结合自注意力机制提升空间关系建模能力	1) 存在模式崩溃风险 2) 训练过程不稳定需精细调参 3) 生成图像可能包含伪影	COVID-19筛查 数据增强 病理特征合成
扩散模型	1) 生成质量达到诊断级精度 2) 支持渐进式病理特征控制 3) 结合语言 - 图像多模态生成(如Roentgen) 4) 在图像去噪任务下鲁棒性强	1) 计算资源需求高 2) 采样速度较慢 3) 需要大规模标注数据支持	高分辨率病理合成 报告 - 图像联合生成 多标签疾病生成

总的来说，在胸部 X 射线生成与编辑领域，VAE 模型及其变体更适合于数据增强和潜在空间分析，但在生成高质量图像方面存在局限，适用于资源有限但需要稳定性的场景；GAN 模型及其变体能够生成细节丰富的图像，但在训练稳定性和伪影控制上存在挑战，适合对图像质量要求较高的任务。扩散模型及其变体在图像质量和多样性上表现出色，但高计算成本和慢采样速度使其更适合资源充足且对保真度要求极高的应用。

7. 研究展望

生成模型在胸部 X 射线图像生成与编辑领域已取得显著的进展，但其在临床转化方面仍然面临着技术瓶颈与跨学科挑战。基于现有研究成果与领域发展的趋势，本文提出以下未来研究的方向：

- 1) 多模态生成与跨域知识融合
- 当前研究主要聚焦于单模态图像合成，亟需构建跨模态生成框架，实现 CXR 与 CT、MRI、超声等多模态影像的协同生成与特征对齐。例如，可借鉴潜在扩散模型(LDM)的跨模态迁移能力，通过解剖结构约束来增强生成图像的临床一致性。近期研究表明，结合区域交叉注意力机制的区域对比描述生成器(如 CoCa-CXR [57])，可以有效捕捉影像对之间的局部差异与时序演变特征，为动态病理建模提供新思路。
- 2) 隐私保护与分布式生成系统
- 医疗数据隐私问题仍是限制生成模型规模化应用的核心障碍。需开发融合差分隐私(DP)与联邦学习

的分布式生成框架，例如基于联邦扩散模型的联合训练范式，在保障数据隔离的前提下实现跨机构知识共享。此外，需建立合成数据的标准化脱敏流程，明确生成影像的元数据标注规则(如 DICOM 标签重建精度)，以满足临床诊断的合规性要求。

3) 病理特征解耦与可控生成

现有模型对复杂病理特征的解耦能力不足，导致生成图像的病变形态与解剖背景缺乏生物学合理性。建议引入因果推理框架，结合解剖标志点约束，实现肺叶级病理特征的可控编辑。同时，可借鉴对比式描述生成器的时序建模能力，通过疾病进程模拟(如肺炎吸收、肺纤维化演变)构建动态病理图谱，为医学教育与罕见病研究提供数据支撑。

4) 模型泛化性与鲁棒性增强

当前模型在跨中心、跨设备数据上的泛化性能显著受限。需探索基于域泛化(Domain Generalization)的迁移学习策略，例如通过对抗性特征对齐消除设备厂商差异(如 GE vs. Siemens X 光机成像特性)。此外，需开发针对生成模型的对抗攻击防御机制，确保合成数据在对抗样本干扰下的稳定性。

5) 伦理规范与临床转化路径

生成式医学影像的伦理风险亟待系统性评估。建议参考 FDA 《AI/ML-Based Synthetic Medical Data》指南，建立合成数据的分级管理制度：① 教育级(允许形态变异)；② 诊断级(需验证解剖保真度)；③ 治疗级(需通过 III 类医疗器械认证)。同时，需明确生成模型在医疗事故中的责任归属，探索“医生-AI”协同诊断的法律边界。

6) 临床集成与循证医学验证

生成模型的终极价值需通过多中心临床试验验证。重点方向包括：① 合成数据对罕见病诊断模型(如粟粒性肺结核检测)的增强效能量化；② 反事实生成在放射科住院医师培训中的应用效果评估；③ 生成式报告系统(如 KARA-CXR)对诊断效率的提升幅度。需建立涵盖影像组学特征、病理金标准与长期随访数据的多维评估体系，推动技术从实验室向床旁转化。

总之，生成模型正重塑胸部 X 射线影像分析的科研范式与临床实践。通过多学科协同创新(如计算机视觉、临床医学、生物伦理学)，未来五年有望实现三大突破：① 跨模态生成框架支撑的全息数字孪生构建；② 隐私安全且合规的分布式医学数据生态；③ 通过 FDA 认证的生成式辅助诊断系统。这些进展将加速精准放射医学从“经验驱动”向“数据智能驱动”的范式转型。

基金项目

伊犁师范大学科研重点项目(2020YSZD004)；伊犁师范大学提升学科综合实力专项自科重点项目(22XKZZ19)。

参考文献

- [1] Moris, D.I., Moura, J.d., Novo, J. and Ortega, M. (2024) Adapted Generative Latent Diffusion Models for Accurate Pathological Analysis in Chest X-Ray Images. *Medical & Biological Engineering & Computing*, **62**, 2189-2212. <https://doi.org/10.1007/s11517-024-03056-5>
- [2] Kingma, D.P. and Welling, M. (2014) Auto-Encoding Variational Bayes. *Stat*, 1050: 1.
- [3] Rais, K., Amroune, M., Benmachiche, A., *et al.* (2024) Exploring Variational Autoencoders for Medical Image Generation: A Comprehensive Study.
- [4] Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., *et al.* (2014) Generative Adversarial Nets. *Communications of the ACM*, **63**, 139-144.
- [5] Alamir, M. and Alghamdi, M. (2022) The Role of Generative Adversarial Network in Medical Image Analysis: An In-Depth Survey. *ACM Computing Surveys*, **55**, 1-36. <https://doi.org/10.1145/3527849>

- [6] Tang, Y., Tang, Y., Zhu, Y., Xiao, J. and Summers, R.M. (2021) A Disentangled Generative Model for Disease Decomposition in Chest X-Rays via Normal Image Synthesis. *Medical Image Analysis*, **67**, Article ID: 101839. <https://doi.org/10.1016/j.media.2020.101839>
- [7] Sohl-Dickstein, J., Weiss, E., Maheswaranathan, N., *et al.* (2015) Deep Unsupervised Learning Using Nonequilibrium Thermodynamics. *International Conference on Machine Learning*, Lille, 6-11 July 2015, 2256-2265.
- [8] Ho, J., Jain, A. and Abbeel, P. (2020) Denoising Diffusion Probabilistic Models. *Proceedings of the 34th International Conference on Neural Information Processing Systems*, Vancouver, 6-12 December 2020, 6840-6851.
- [9] Rombach, R., Blattmann, A., Lorenz, D., Esser, P. and Ommer, B. (2022) High-Resolution Image Synthesis with Latent Diffusion Models. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, 18-24 June 2022, 10684-10695. <https://doi.org/10.1109/cvpr52688.2022.01042>
- [10] Kazerouni, A., Aghdam, E.K., Heidari, M., Azad, R., Fayyaz, M., Hachililoglu, I., *et al.* (2023) Diffusion Models in Medical Imaging: A Comprehensive Survey. *Medical Image Analysis*, **88**, Article ID: 102846. <https://doi.org/10.1016/j.media.2023.102846>
- [11] Shin, H., Tenenholtz, N.A., Rogers, J.K., Schwarz, C.G., Senjem, M.L., Gunter, J.L., *et al.* (2018) Medical Image Synthesis for Data Augmentation and Anonymization Using Generative Adversarial Networks. *Simulation and Synthesis in Medical Imaging: 3rd International Workshop, SASHIMI 2018, Held in Conjunction with MICCAI 2018*, Granada, 16 September 2018, 1-11. https://doi.org/10.1007/978-3-030-00536-8_1
- [12] Chambon, P., Bluethgen, C., Delbrouck, J.B., *et al.* (2022) Roentgen: Vision-Language Foundation Model for Chest X-Ray Generation.
- [13] Lee, H., Lee, D.Y., Kim, W., *et al.* (2023) Vision-Language Generative Model for View-Specific Chest X-Ray Generation.
- [14] Yi, X., Walia, E. and Babyn, P. (2019) Generative Adversarial Network in Medical Imaging: A Review. *Medical Image Analysis*, **58**, Article ID: 101552. <https://doi.org/10.1016/j.media.2019.101552>
- [15] Pu, Y., Gan, Z., Henao, R., *et al.* (2016) Variational Autoencoder for Deep Learning of Images, Labels and Captions. *30th Conference on Neural Information Processing Systems (NIPS 2016)*, Barcelona, 5-10 December 2016.
- [16] Karras, T., Laine, S. and Aila, T. (2019) A Style-Based Generator Architecture for Generative Adversarial Networks. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, 15-20 June 2019, 4401-4410. <https://doi.org/10.1109/cvpr.2019.00453>
- [17] Song, Y., Sohl-Dickstein, J., Kingma, D.P., *et al.* (2020) Score-Based Generative Modeling through Stochastic Differential Equations.
- [18] Park, E. (2015) Manifold Learning with Variational Auto-Encoder for Medical Image Analysis. Technical Report, University of North Carolina, Tech. Rep.
- [19] Bercea, C.I., Rueckert, D. and Schnabel, J.A. (2022) What Do We Learn? Debunking the Myth of Unsupervised Outlier Detection.
- [20] Burgess, C.P., Higgins, I., Pal, A., *et al.* (2018) Understanding Disentangling in beta-VAE.
- [21] Zhou, L., Deng, W. and Wu, X. (2020) Unsupervised Anomaly Localization Using VAE and beta-VAE.
- [22] Ghali, R. and Akhloufi, M.A. (2023) Vision Transformers for Lung Segmentation on CXR Images. *SN Computer Science*, **4**, Article No. 414. <https://doi.org/10.1007/s42979-023-01848-4>
- [23] Crespi, L., Loiacono, D. and Chiti, A. (2021) Chest X-Rays Image Classification from β -Variational Autoencoders Latent Features. *2021 IEEE Symposium Series on Computational Intelligence (SSCI)*, Orlando, 5-7 December 2021, 1-8. <https://doi.org/10.1109/ssci50451.2021.9660190>
- [24] Cao, F. and Zhao, H. (2021) Automatic Lung Segmentation Algorithm on Chest X-Ray Images Based on Fusion Variational Auto-Encoder and Three-Terminal Attention Mechanism. *Symmetry*, **13**, Article No. 814. <https://doi.org/10.3390/sym13050814>
- [25] Gerlings, J., Jensen, M.S. and Shollo, A. (2021) Explainable AI, but Explainable to Whom? An Exploratory Case Study of xAI in Healthcare. In: Lim, C.-P., *et al.*, Eds., *Intelligent Systems Reference Library*, Springer International Publishing, 169-198. https://doi.org/10.1007/978-3-030-83620-7_7
- [26] Chatterjee, S., Maity, S., Bhattacharjee, M., Banerjee, S., Das, A.K. and Ding, W. (2022) Variational Autoencoder Based Imbalanced COVID-19 Detection Using Chest X-Ray Images. *New Generation Computing*, **41**, 25-60. <https://doi.org/10.1007/s00354-022-00194-y>
- [27] Montenegro, H. and Cardoso, J.S. (2024) Anonymizing Medical Case-Based Explanations through Disentanglement. *Medical Image Analysis*, **95**, Article ID: 103209. <https://doi.org/10.1016/j.media.2024.103209>
- [28] Gu, Y., Yang, J., Usuyama, N., *et al.* (2023) Biomedjourney: Counterfactual Biomedical Image Generation by Instruction-Learning from Multimodal Patient Journeys.

- [29] Li, F., Huang, W., Luo, M., Zhang, P. and Zha, Y. (2021) A New VAE-GAN Model to Synthesize Arterial Spin Labeling Images from Structural MRI. *Displays*, **70**, Article ID: 102079. <https://doi.org/10.1016/j.displa.2021.102079>
- [30] Siddiqui, A.A., Tirunagari, S., Zia, T., *et al.* (2024) VALD-MD: Visual Attribution via Latent Diffusion for Medical Diagnostics.
- [31] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., *et al.* (2020) Generative Adversarial Networks. *Communications of the ACM*, **63**, 139-144. <https://doi.org/10.1145/3422622>
- [32] 刘建伟, 谢浩杰, 罗雄麟. 生成对抗网络在各领域应用研究进展[J]. 自动化学报, 2020, 46(12): 2500-2536.
- [33] Aljohani, A. and Alharbe, N. (2022) Generating Synthetic Images for Healthcare with Novel Deep Pix2Pix GAN. *Electronics*, **11**, Article No. 3470. <https://doi.org/10.3390/electronics11213470>
- [34] Kora Venu, S. and Ravula, S. (2020) Evaluation of Deep Convolutional Generative Adversarial Networks for Data Augmentation of Chest X-Ray Images. *Future Internet*, **13**, Article No. 8. <https://doi.org/10.3390/fi13010008>
- [35] Dhawan, K. and Nijhawan, S.S. (2024) Cross-Modality Synthetic Data Augmentation Using GANs: Enhancing Brain MRI and Chest X-Ray Classification.
- [36] Huang, Y., Maier, A., Fan, F., *et al.* (2022) Learning Perspective Deformation in X-Ray Transmission Imaging.
- [37] Liang, Z., Huang, J.X. and Antani, S. (2022) Image Translation by Ad CycleGAN for COVID-19 X-Ray Images: A New Approach for Controllable Gan. *Sensors*, **22**, Article No. 9628. <https://doi.org/10.3390/s22249628>
- [38] Kong, L., Lian, C., Huang, D., *et al.* (2021) Breaking the Dilemma of Medical Image-to-Image Translation. *Advances in Neural Information Processing Systems*, Vol. 34, 1964-1978.
- [39] Qin, X., Bui, F.M., Han, Z. and Khademi, A. (2024) Toward Improved Interpretability in Medical Imaging: Revealing the Disease Evidence from Chest X-Ray Images Using an Adversarial Generative Approach. *IEEE Access*, **12**, 82002-82014. <https://doi.org/10.1109/access.2024.3412608>
- [40] Kim, E., Lee, S. and Lee, K.M. (2023) Abnormality Detection in Chest X-Ray via Residual-Saliency from Normal Generation. *IEEE Access*, **11**, 21799-21810. <https://doi.org/10.1109/access.2023.3251350>
- [41] Saboo, A., Ramachandran, S.N., Dierkes, K., *et al.* (2021) Towards Disease-Aware Image Editing of Chest X-Rays.
- [42] Jin, Y., Chang, W. and Ko, B. (2023) Generating Chest X-Ray Progression of Pneumonia Using Conditional Cycle Generative Adversarial Networks. *IEEE Access*, **11**, 88152-88160. <https://doi.org/10.1109/access.2023.3305994>
- [43] Liang, Z., Huang, J.X., Li, J. and Chan, S. (2020). Enhancing Automated COVID-19 Chest X-Ray Diagnosis by Image-To-Image GAN Translation. 2020 *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Seoul, 16-19 December 2020, 1068-1071. <https://doi.org/10.1109/bibm49941.2020.9313466>
- [44] Weber, T., Ingrisch, M., Bischl, B. and Rügamer, D. (2023) Implicit Embeddings via GAN Inversion for High Resolution Chest Radiographs. In: Fragemann, J., *et al.*, Eds., *Medical Applications with Disentanglements*, Springer, 22-32. https://doi.org/10.1007/978-3-031-25046-0_3
- [45] Saboo, A., Gyawali, P.K., Shukla, A., *et al.* (2021) Latent-Optimization Based Disease-Aware Image Editing for Medical Image Augmentation. *BMVC*, 22-25 November 2021, 181.
- [46] Wang, Z., Zhang, X., Chen, W. and Niu, J. (2022) Lung Segmentation Reconstruction Based Data Augmentation Approach for Abnormal Chest X-Ray Images Diagnosis. 2022 *44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Glasgow, 11-15 July 2022, 2203-2207. <https://doi.org/10.1109/embc48229.2022.9871784>
- [47] Irvin, J., Rajpurkar, P., Ko, M., Yu, Y., Ciurea-Ilcus, S., Chute, C., *et al.* (2019) Chexpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison. *Proceedings of the AAAI Conference on Artificial Intelligence*, **33**, 590-597. <https://doi.org/10.1609/aaai.v33i01.3301590>
- [48] Khader, F., Müller-Franzes, G., Tayebi Arasteh, S., Han, T., Haarbuerger, C., Schulze-Hagen, M., *et al.* (2023) Denoising Diffusion Probabilistic Models for 3D Medical Image Generation. *Scientific Reports*, **13**, Article No. 7303. <https://doi.org/10.1038/s41598-023-34341-2>
- [49] Schaudt, D., Späte, C., von Schwerin, R., Reichert, M., von Schwerin, M., Beer, M., *et al.* (2023) A Critical Assessment of Generative Models for Synthetic Data Augmentation on Limited Pneumonia X-Ray Data. *Bioengineering*, **10**, Article No. 1421. <https://doi.org/10.3390/bioengineering10121421>
- [50] Packhäuser, K., Folle, L., Thamm, F. and Maier, A. (2023). Generation of Anonymous Chest Radiographs Using Latent Diffusion Models for Training Thoracic Abnormality Classification Systems. 2023 *IEEE 20th International Symposium on Biomedical Imaging (ISBI)*, Cartagena, 18-21 April 2023, 1-5. <https://doi.org/10.1109/isbi53787.2023.10230346>
- [51] Dhariwal, P. and Nichol, A. (2021) Diffusion Models Beat Gans on Image Synthesis. *Advances in Neural Information Processing Systems*, Vol. 34, 8780-8794.

-
- [52] Weber, T., Ingrisch, M., Bischl, B. and Rügamer, D. (2023) Cascaded Latent Diffusion Models for High-Resolution Chest X-Ray Synthesis. In: *Lecture Notes in Computer Science*, Springer, 180-191.
https://doi.org/10.1007/978-3-031-33380-4_14
 - [53] Wolleb, J., Bieder, F., Sandkühler, R. and Cattin, P.C. (2022) Diffusion Models for Medical Anomaly Detection. In: Kashima, H., Ide, T. and Peng, W.-C., Eds., *Advances in Knowledge Discovery and Data Mining*, Springer, 35-45.
https://doi.org/10.1007/978-3-031-16452-1_4
 - [54] Fathi, N., Kumar, A., Nichyporuk, B., *et al.* (2024) DeCoDEX: Confounder Detector Guidance for Improved Diffusion-based Counterfactual Explanations.
 - [55] Hashmi, A.U.R., Almakky, I., Qazi, M.A., *et al.* (2024) Xreal: Realistic Anatomy and Pathology-Aware X-Ray Generation via Controllable Diffusion Model.
 - [56] Song, J., Meng, C. and Ermon, S. (2020) Denoising Diffusion Implicit Models.
 - [57] Chen, Y., Xu, S., Sellergren, A., *et al.* (2025) CoCa-CXR: Contrastive Captioners Learn Strong Temporal Structures for Chest X-Ray Vision-Language Understanding.