

Unsupervised Medical Image Translation with Adversarial Diffusion Models

摘要

通过源-目标模态转换对缺失图像进行补全（也称图像翻译）可以丰富医学图像模态的多样性，有助于降低成本、提高效率、提高临床诊断准确度。利用传统的 GAN 进行一次推理，隐式地表征图像分布，可能会使模型训练效果受到样本保真度的限制。作者提出了一种基于对抗扩散建模的新方法 SynDiff，以提高医学图像翻译的性能。为了捕获图像分布的直接相关性，SynDiff 利用条件扩散过程，逐步将噪声和源图像映射到目标图像上。为了在推理过程中快速准确地进行图像采样，在反向扩散过程中采用了一种对抗投影模块实现加大步长扩散，有效提高了推理效率。同时作者设计了一个循环一致架构，使得网络可以在不配对的数据集上进行训练，该架构结合了扩散和非扩散模块，可以在提高效率的同时保证生成图像的保真度。SynDiff 在多对比 MRI 和 MRI-CT 翻译任务中对比 GAN 和其他扩散模型有着显著优势。

关键词：医学图像翻译；GAN；扩散模型；加大步长扩散；循环一致

1 引言

多模态成像在评估人体解剖和功能中十分关键，但由于成本原因，临床中广泛采用多模式扫描不切实际。GAN 模型的隐式表征方式可能会导致学习偏差，包括过早收敛和模式崩溃。此外，GAN 模型采用快速的单次采样过程，没有中间步骤，限制了映射的可靠性。最近的计算机视觉研究采用了基于显式似然表征和渐进采样过程的扩散模型，以提高无条件生成建模任务中的样本保真度 [?], [1]。作者提出了一种新的对抗扩散模型 SynDiff，用于医学图像合成，以执行高效和高保真的模态转换 (图 2)。与常规扩散模型不同，SynDiff 采用了快速扩散过程和大步长来提高效率。在反向扩散步骤中，通过一种新的源条件对抗式投影模块实现精确采样，该投影模块在源图像的引导下对目标图像样本进行降噪。为了实现无监督学习，设计了一个循环一致的架构，在两种模式之间双边耦合扩散和非扩散过程 (图 3)，对多对比 MRI 和 MRI-CT 翻译进行了全面的演示。实验结果清楚地表明 SynDiff 相对于 GAN 和其他扩散模型的优越性。

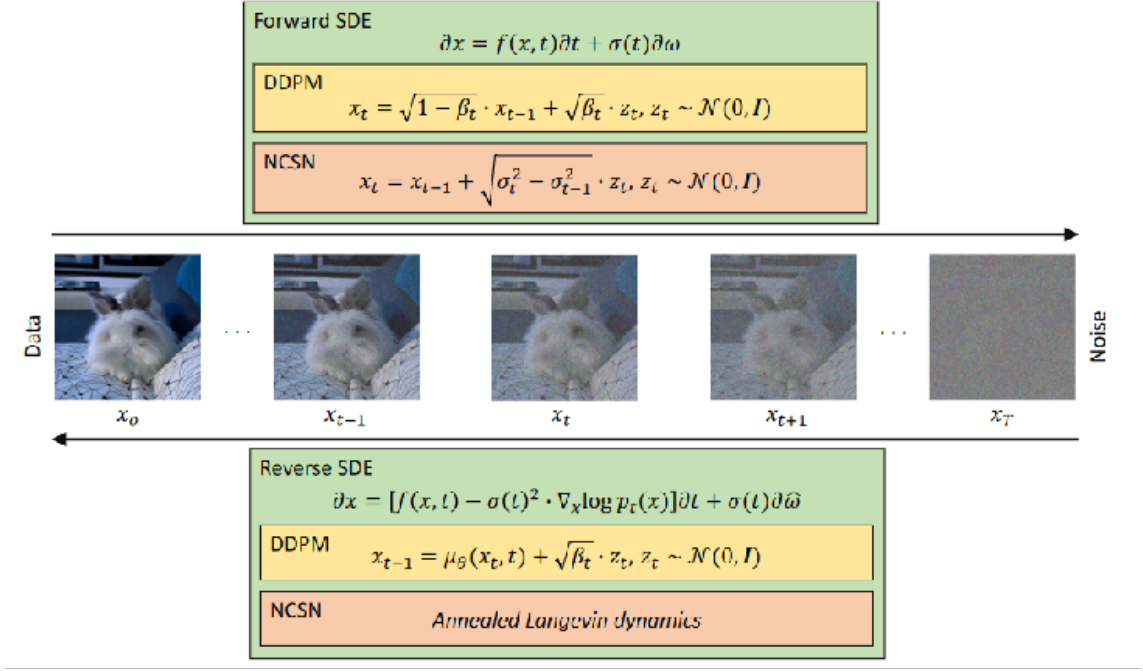


图 1. Diffusion Model 架构

2 相关工作

2.1 Diffusion Model

作为 GAN 的最新替代品，深度扩散模型（图 1）在计算机视觉的生成建模任务中受到了关注 [?], [1]。扩散模型从纯噪声样本开始，通过重复去噪，从期望的分布中生成图像样本。基于扩散的方法最近被用于单峰成像任务，如图像重建 [2]- [3]、无条件图像生成 [4] 和异常检测 [5]、[6]。然而，这些方法通常只能处理单模态图像。此外，目前的方法通常推理步骤过多，耗时过长。

2.2 近期相关研究

自然图像生成模型 DDGAN 试图提高逆向扩散步骤的效率 [7]。DDGAN 使用一个对抗发生器来进行反向扩散，而不需要源图像的引导。相比之下，SynDiff 是基于有条件的扩散过程，在源图像和目标图像之间进行转换。除了扩散模块，SynDiff 还包含了非扩散模块来实现无监督的图像翻译。一项关于自然图像无监督翻译的研究提出了一种非对抗扩散模型 UNIT-DDPM [8]。基于源-目标模式共享潜在空间的概念，UNIT-DDPM 使用并行扩散过程在大量反向步骤中同时为两种模式生成样本；从源扩散过程中提取的噪声源图像样本用于目标扩散过程中目标图像的生成。SynDiff 使用对抗性投影模块在几个步骤中进行有效采样。最近的一项研究独立考虑了一种基于条件评分的方法 UMM-CGSM，用于在多对比 MRI 方案中插入缺失序列 [9]。UMM-CGSM 推理步骤相对较多，并对成对数据集进行监督训练。SynDiff 采用了一种对抗扩散模型，在几步内进行有效的采样，还可以进行无监督学习。

3 作者方法

3.1 对抗投影模块

与常规的扩散模型较小的步长、较多的扩散步骤相比，本文引入了一个对抗投影模块（图2），其目的是加大步长扩散、提高扩散效率；对抗投影模块利用了传统 GAN 的生成器和对抗器架构，在加大步长这一操作会打破扩散模型依赖的高斯假设的前提下，很好地桥接了反向扩散过程中的每一小步。

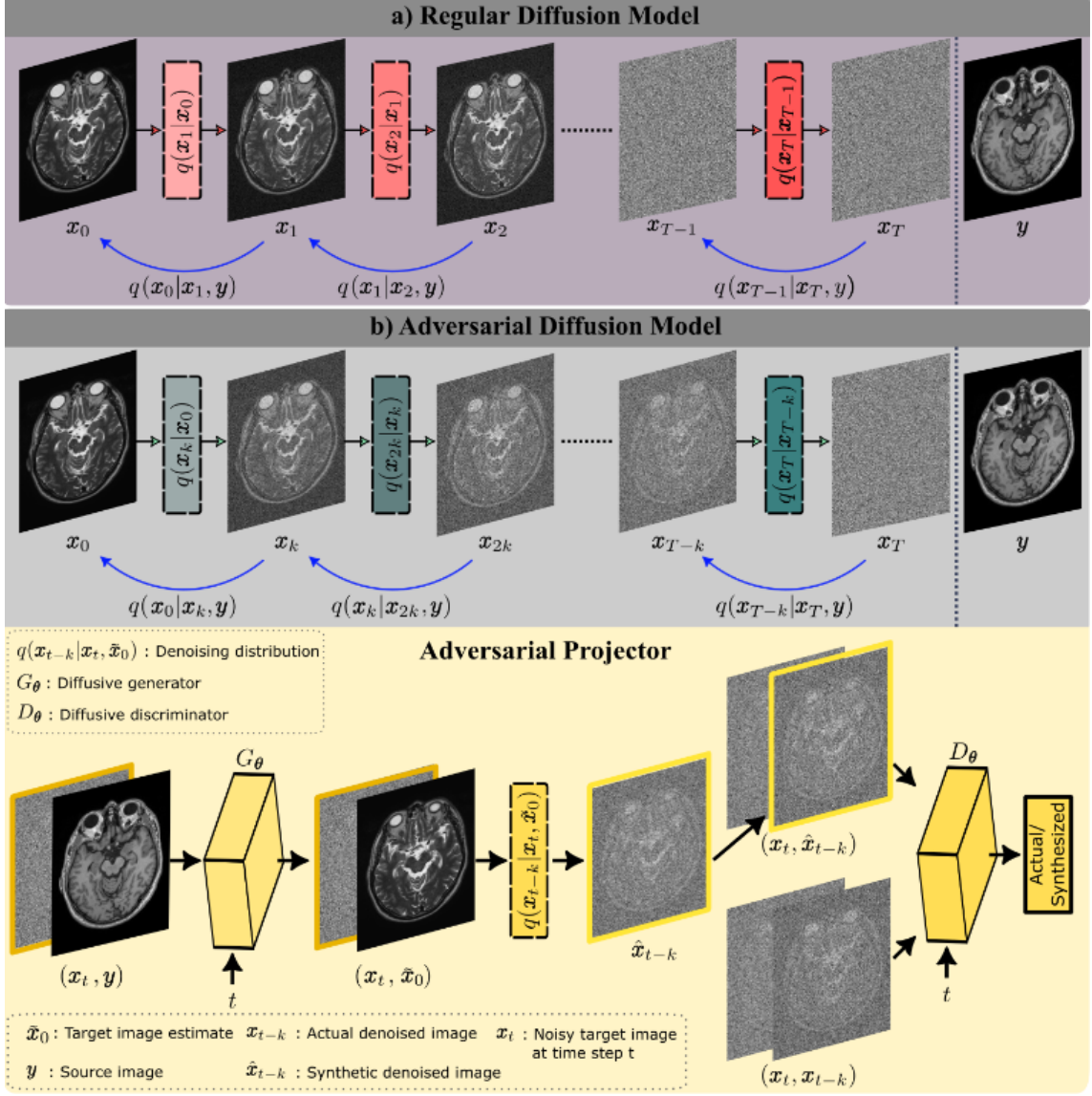


图 2. 对抗投影模块

3.2 循环一致模块

为了从未配对的训练集中学习，作者引入了一个基于非扩散和扩散模块的循环一致架构（图 3），该架构在两种模态之间双向转换。SynDiff 利用非扩散模块来估计与训练集中每个目标图像配对的源图像，作为训练的引导。在给定模态 A 的目标图像的情况下，产生模态 B 的源图像估计；在给定目标图像的情况下，产生源图像估计。扩散模块以来自非扩散模块的源图

像估计作为指导，以合成目标图像。

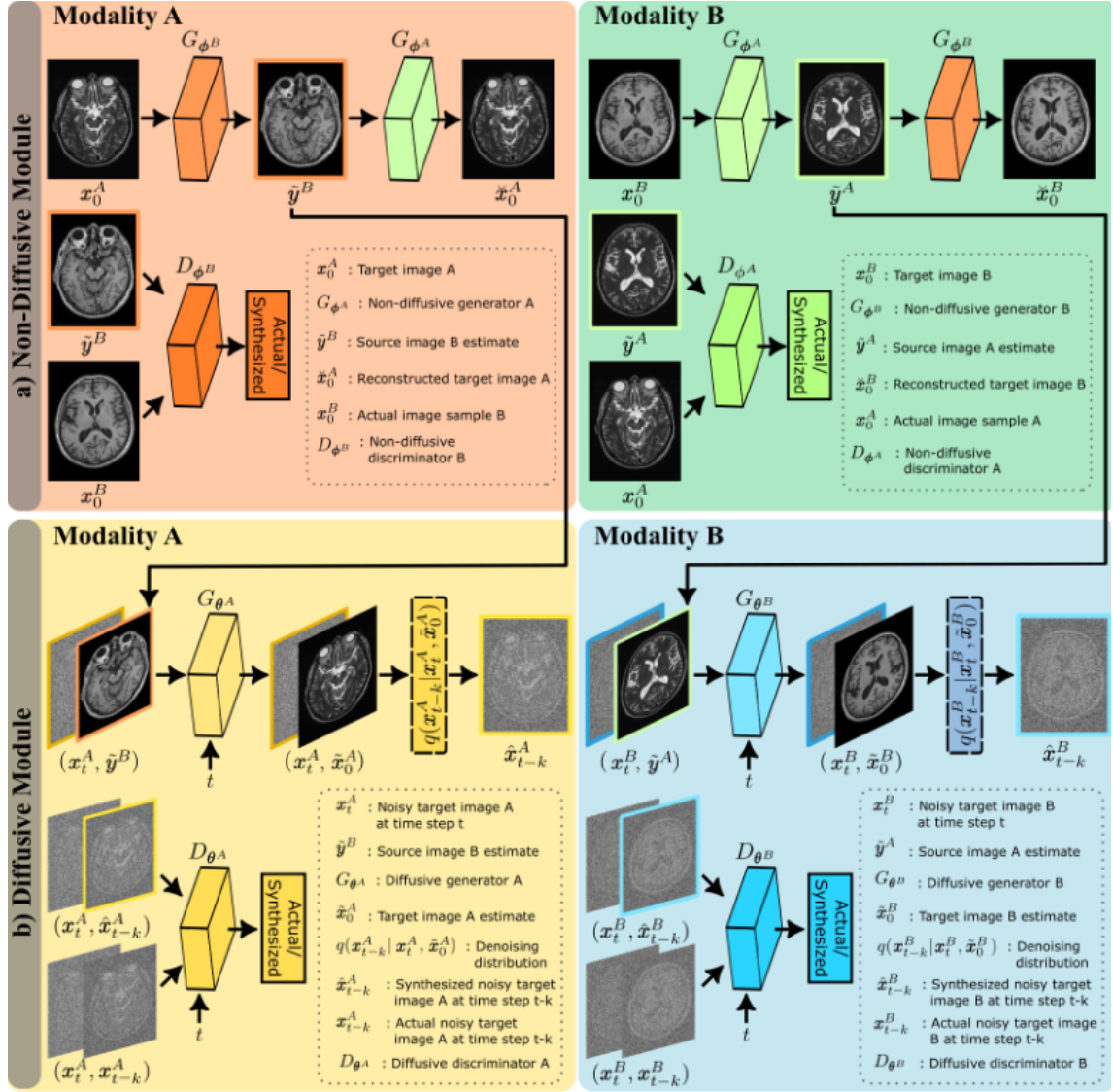


图 3. 耦合非扩散-扩散模块

4 损失函数

循环一致性 loss 定义为:

$$L_{cyc} = E_{t, q(x_0^{A,B}), q(x_t^{A,B} | x_0^{A,B})} [\lambda_{1\phi} (|x_0^A - \check{x}_0^A|_1 + |x_0^B - \check{x}_0^B|_1) + \lambda_{1\theta} (|x_0^A - \check{x}_0^A|_1 + |x_0^B - \check{x}_0^B|_1)] \quad (1)$$

总的生成器 loss 为:

$$L_G^{total} = \lambda_{2\phi} (L_{G_\phi^A} + L_{G_\phi^B}) + \lambda_{2\theta} (L_{G_\theta^A} + L_{G_\theta^B}) + L_{cyc} \quad (2)$$

总的鉴别器 loss 为:

$$L_D^{total} = \lambda_{2\phi} (L_{D_\phi^A} + L_{D_\phi^B}) + \lambda_{2\theta} (L_{D_\theta^A} + L_{D_\theta^B}) \quad (3)$$

5 复现细节

5.1 数据集

原文中使用了三个数据集。使用了其中两个公开的 MRI 数据集进行复现，分别是：IXI 脑部健康 MRI 图像，BRATS 脑部肿瘤 MRI 扫描图像。两个数据集都包含 T1, T2 和 PD 三种模态，复现了一个数据集内不同模态的相互转换。

5.2 与已有开源代码对比

参考了原作者提供的代码，但其中并未包括数据预处理、指标计算部分；自己撰写了数据预处理部分，包括将训练数据配对、分组，取（训练，验证，测试）集，将其三维图像切片及转换为.mat 文件输入网络；源代码实现分布式训练代码有 bug，在其基础上进行改进调整；撰写指标计算评估部分。

5.3 实验环境搭建

```
python>=3.6.9  
torch>=1.7.1  
torchvision>=0.8.2  
cuda=>11.2  
ninja  
python3.x-dev
```

6 实验结果分析

对于脑部健康数据集的 MRI 模态 T1 和 PD 之间的翻译展示结果如下。（左：源模态，右：目标模态）

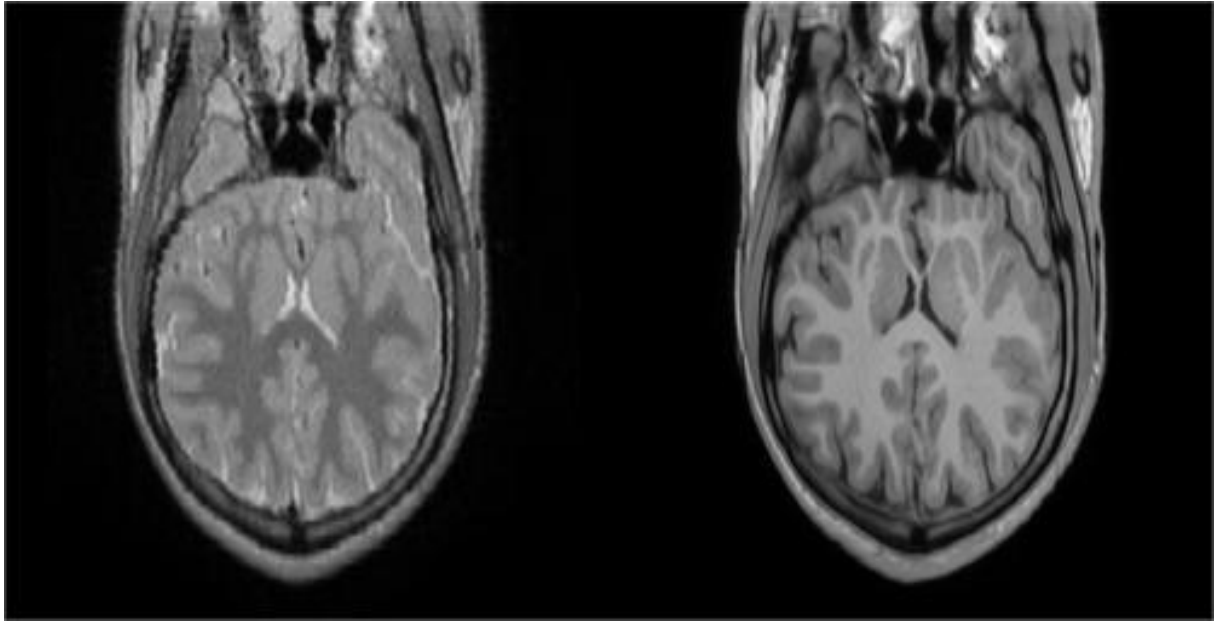


图 4. T1-PD 测试结果 (IXI)

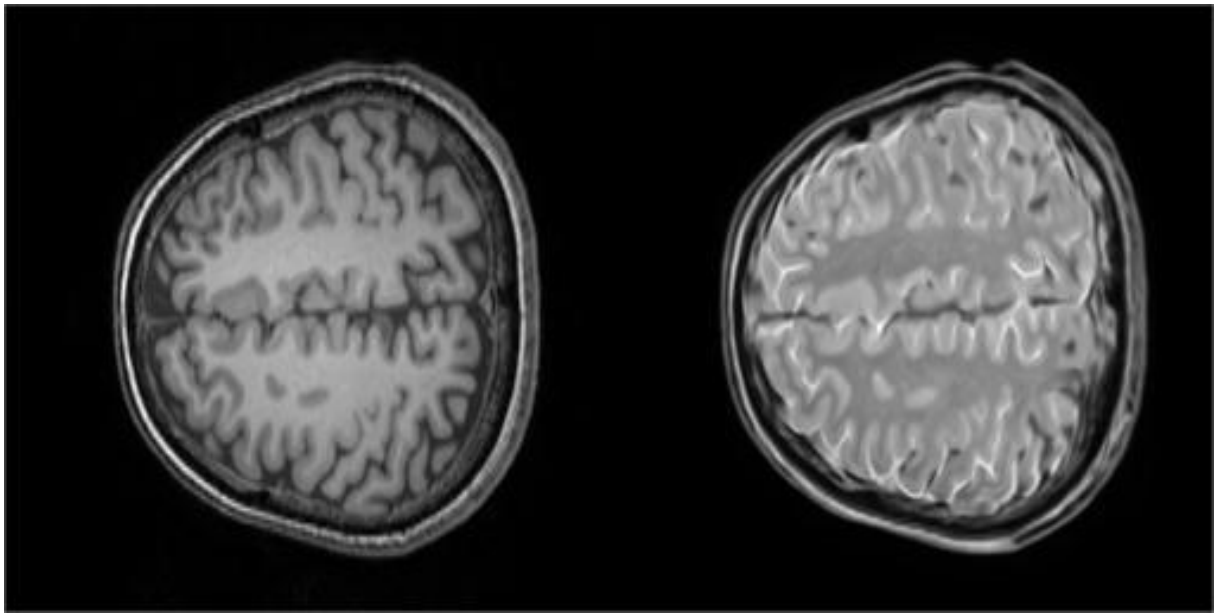


图 5. PD-T1 测试结果 (IXI)

FID 展示了源模态图像和目标模态图像之间的差异。复现工作与原文 FID 指标对比：

	$T_1(IXI)$	$T_2(BRATS)$
原文	30.75	75.04
复现	31.05	77.56

表 1. 复现结果与论文对比：FID

7 总结与展望

在这项研究中，作者介绍了一种新的对抗扩散模型，用于医学图像在源和目标模式之间的转换。SynDiff 利用快速扩散过程有效地合成目标图像，并使用条件对抗投影仪进行准确的储备扩散采样。通过循环一致的架构实现无监督学习，该架构体现了两种模式之间的耦合扩散过程。与最先进的 GAN 和扩散模型相比，SynDiff 实现了卓越的质量，并且在高保真医学图像翻译方面具有很大的前景。

然而，文章中并未提到对于预训练的模型在其他数据集或模态上的表现，因此没有展示出模型的通用性；评估指标一项中缺乏专家评估环节，单凭自然图像处理指标不一定能很好地使模型运用到临床上辅助医生诊断。未来的工作可以从以下几个方面进行展望：

1. 提高模型的通用性：可以通过提高数据集的多样性，以及模型对不同模态图像的深层特征的学习能力，提高模型处理不同输入图像时的稳定性，以及模型的可靠性。
2. 引入新的评估指标：可以引入专家为测试结果打分，作为评估模型有效性的有力指标。
3. 扩展到其他医学图像任务：扩散模型作为一种性能优秀的生成模型，可以作为上游任务为其它任务奠定重要的基础。未来可以将其结果用于其他方向任务，如病灶分割、病症分类等。

总的来说，虽然作者的工作存在一些不足，但它为我们提供了一个重要的研究方向，期待未来的研究能够在这个方向上取得更多的进展。

参考文献

- [1] J. Ho, A. Jain, and P. Abbeel. Denoising diffusion probabilistic models. In *Adv Neural Inf Process Syst*, volume 33, pages 6840–6851, 2020.
- [2] A. Jalal, M. Arvinte, G. Daras, E. Price, A. G. Dimakis, and J. Tamir. Robust compressed sensing mri with deep generative priors. In *Adv Neural Inf Process Syst*, volume 34, pages 14938–14954, 2021.
- [3] H. Chung, E. S. Lee, and J. C. Ye. Mr image denoising and superresolution using regularized reverse diffusion. *arXiv preprint arXiv:2203.12621*, 2022.
- [4] W. H. L. Pinaya, P.-D. Tudosiu, J. Dafflon, P. F. da Costa, V. Fernandez, P. Nachev, et al. Brain imaging generation with latent diffusion models. *arXiv preprint arXiv:2209.07162*, 2022.
- [5] J. Wolleb, F. Bieder, R. Sandkuhler, and P. C. Cattin. Diffusion models for medical anomaly detection. In *Med Image Comput Comput Assist Inter*, volume 13438, pages 35–45. Springer, 2022.
- [6] W. H. L. Pinaya, M. S. Graham, R. Gray, P. F. Da Costa, P.-D. Tudosiu, P. Wright, et al. Fast unsupervised brain anomaly detection and segmentation with diffusion models. *arXiv preprint arXiv:2206.03461*, 2022.

- [7] Z. Xiao, K. Kreis, and A. Vahdat. Tackling the generative learning trilemma with denoising diffusion gans. In *Int Conf Learn Represent*, 2022.
- [8] H. Sasaki, C. G. Willcocks, and T. P. Breckon. Unit-ddpm: Unpaired image translation with denoising diffusion probabilistic models. *arXiv preprint arXiv:2104.05358*, 2021.
- [9] X. Meng, Y. Gu, Y. Pan, N. Wang, P. Xue, M. Lu, et al. A novel unified conditional score-based generative framework for multi-modal medical image completion. *arXiv preprint arXiv:2207.03430*, 2022.