Three-Dimensional Medical Image Synthesis with Denoising Diffusion Probabilistic Models

Zolnamar Dorjsembe[1] ZOLNAMAR@GMAIL.COM, M10915809@MAIL.NTUST.EDU.TW
Sodtavilan Odonchimed[2] 1184142454@G.ECC.U-TOKYO.AC.JP
Furen Xiao[3] XIAO@NTUH.GOV.TW
[1] *Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology, Taipei 106, Taiwan*
[2] *Faculty of Engineering, The University of Tokyo, Tokyo 113-8654, Japan*
[3] *Department of Neurosurgery, National Taiwan University Hospital, Taipei 100, Taiwan*

Editors: Under Review for MIDL 2022

Abstract

Denoising diffusion probabilistic models (DDPM) have recently shown superior performance in image synthesis and have been extensively studied in various image processing tasks. In this work, we propose a 3D-DDPM for generating three-dimensional (3D) medical images. Different from previous studies, to the best of our knowledge, this work presents the first attempt to investigate the DDPM to enable 3D medical image synthesis. Our study examined the generation of high-resolution magnetic resonance images (MRI) of brain tumors. The proposed method is evaluated through experiments on a semi-public dataset, with both quantitative and qualitative tests showing promising results. Our code will be publicly available at https://github.com/DL-Circle/3D-DDPM.

Keywords: Diffusion models, image synthesis, magnetic resonance imaging (MRI).

1. Introduction

Denoising Diffusion Probabilistic Models (DDPMs) (Ho et al., 2020; Nichol and Dhariwal, 2021) have emerged as a powerful family of generative models that exhibit superior performance and have been extensively studied. In this paper, we present a 3D-DDPM that enables high-quality three-dimensional medical image synthesis. For simplicity, we adapt a structure of DDPM (Ho et al., 2020) with modifications from improved DDPM (Nichol and Dhariwal, 2021) to transform it into the 3D DDPM model. Our model was tested with only 250 timesteps and a simple $L1$ loss. We reproduced the results of state-of-the-art 3D GANs, Cycle Consistent Embedding GAN (CCE-GAN) (Xing et al., 2021) and 3D-α-WGAN (Kwon et al., 2019), and compared our proposed model with them. Our experiments with brain MRI synthesis show that the 3D-DDPM can generate high-quality samples and outperforms the baseline models.

2. Methods

We made several changes to the original DDPM structure. All 2D operations, layers, and noise inputs were replaced by 3D. Structure and formulation of (Ho et al., 2020) are used as our base and the cosine noise schedule proposed in (Nichol and Dhariwal, 2021) is adapted. The forward noising process q for an given input image x_0 by gradually adding small amounts

of Gaussian noise based on the cosine noise schedule in T steps sampled from uniform distribution. The reverse process p_θ is learned by optimizing the model parameters. The main architecture of the proposed method is shown in Figure 1. Considering the space limitations here, we have only provided an overview of the architecture. More details, such as the mathematical formulations, the overall training, and the sampling procedures, can be found in our Github repository [1].

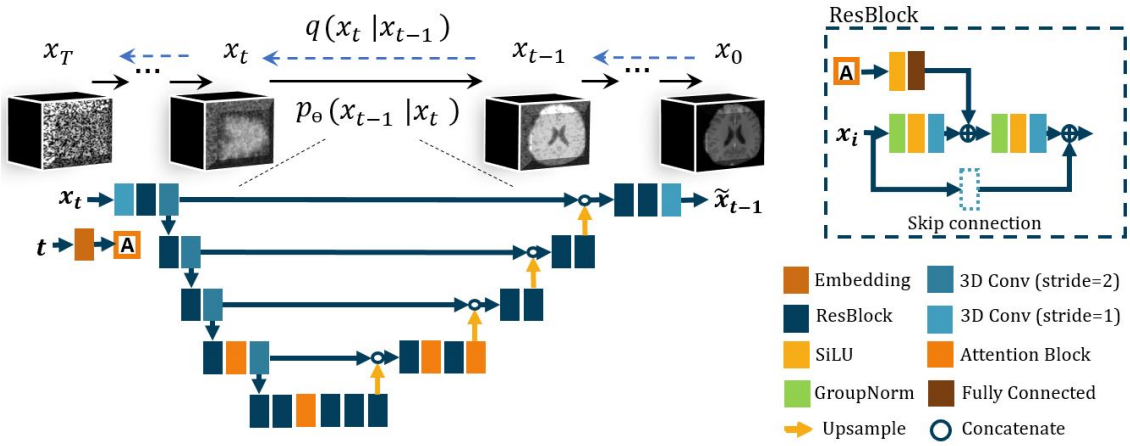

Figure 1: Model architecture

3. Experiments and Results

Dataset: We trained our model on National Taiwan University Hospital's Intracranial Tumor Segmentation (ICTS) dataset[2] which contains 1500 contrast-enhanced anonymized T1 images. We further processed with MRIPreprocessor[3]: for brain extraction, cropped and resized them to be of size 128 x 128 x 128, rescaled the intensities to [-1,1] and applied only horizontal flip augmentation. In addition, we resized the images to 64 $\times$ 64 $\times$ 64 in order to compare results of GAN models.

Models and Training: We compared the following models: (1) 3D-α-WGAN (Kwon et al., 2019), (2) CCE-GAN (Xing et al., 2021). All models are trained on Tesla V100–SXM2 32 GB GPU card with the batch size of 4 for 50K iterations. Our DDPM model is trained with the simple $L1$ loss, the cosine noise schedule for 250 steps, a learning rate of 10^{-4} for the Adam optimizer, number of channels in the first layer is chosen as 64, and used one attention head at resolution 16.

Evaluation: For a quantitative assessment, we chose the pair-wise Multi-Scale Structural Similarity (MS-SSIM)(Kwon et al., 2019), which measures the diversity of generated images. MS-SSIM scores were calculated with the average from 1000 sample pairs for each model. The qualitative evaluation was done by visually assessing the images. We asked two neuroradiologists and a neurosurgeon, each with over 15 years of experience, to classify real and fake MRI images generated by models and the original dataset. Experts were shown 200 images consisting of a 50/50/50/50 mixture of real and synthetic images of our model and 2 baseline models. They were asked to classify it based only on their visual assessment.

1. github.com/DL-Circle/3D-DDPM
2. icts.ml, for more information: https://pubmed.ncbi.nlm.nih.gov/34560272/
3. github.com/ReubenDo/MRIPreprocessor

Results: Table 1 shows the advantages of our model. Baseline CCE-GAN and our model can generate diverse samples with relatively similar scores to that of the real data. Center-cut slices of real and generated samples are shown in Figure 2 and 3. Both baseline models produce blurry results. Surprisingly, for the qualitative assessment, experts classified real images with an average accuracy of 40%. 3D-DDPM images were classified as real with 60%, which shows how strongly our model generates visually accurate images. In contrast, none of the samples generated from baseline models were classified as real.

Table 1: Quantitative results

Model	MS-SSIM
3D-α-WGAN	0.9116
CCE-GAN	0.8238
3D-DDPM (Ours)	0.8241
Real	0.8448

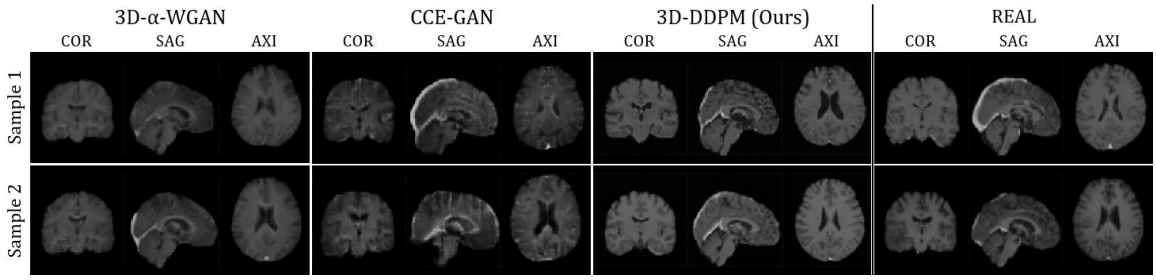

Figure 2: Real and generated samples (size: 64 × 64 × 64)

4. Conclusion

In this paper, we present a novel 3D-DDPM that generates realistic high-quality 3D brain MRI data and was trained on a small set of training images with only 250 steps. As a result, the proposed method outperforms alternative structures in capturing real data distribution and generating diverse samples. One limitation of this study is the use of a small amount of training data. In future work, we will work on other publicly available datasets and extend our model for medical image translation tasks.

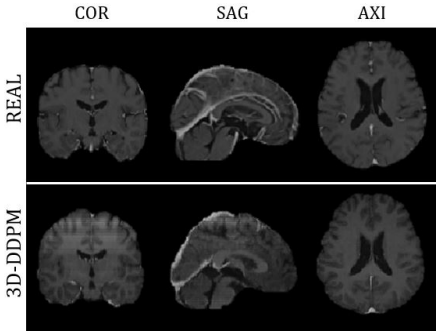

Figure 3: Real and generated sample (size: 128 × 128 × 128)

References

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models, 2020. URL https://arxiv.org/abs/2006.11239.

Gihyun Kwon, Chihye Han, and Dae-shik Kim. Generation of 3d brain mri using auto-encoding generative adversarial networks, 2019. URL https://arxiv.org/abs/1908.02498.

Alex Nichol and Prafulla Dhariwal. Improved denoising diffusion probabilistic models, 2021. URL https://arxiv.org/abs/2102.09672.

Shibo Xing, Harsh Sinha, and Seong Jae Hwang. Cycle consistent embedding of 3d brains with auto-encoding generative adversarial networks. In *Medical Imaging with Deep Learning*, 2021. URL https://openreview.net/forum?id=jgBzGIG-kB.