Medical Imaging Process & Technology (2023) Volume 6 Issue 1
doi: 10.24294/mipt.v6i1.2518

ORIGINAL RESEARCH ARTICLE

2D brain MRI image synthesis based on lightweight denoising diffusion probabilistic model

Jincheng Peng, Guoyue Chen*, Kazuki Saruta, Yuki Terata

Department of Information and Computer Science, Akita Prefectural University, Akita Yurihonjo 015-0055, Japan

*** Corresponding author:** Guoyue Chen, chen@akita-pu.ac.jp

ABSTRACT

In recent years, brain health has received increasing attention, but conventional acquisition of brain MRI (magnetic resonance imaging) images still suffer from issues such as missing data, artifacts, and high costs, which hinders research and diagnosis. With the application of deep learning in medical image synthesis, low-cost, efficient, and high-quality medical MRI synthesis techniques have become a prominent research focus and have gradually matured. However, traditional methods for synthesizing magnetic resonance imaging (MRI) mostly rely on generative adversarial networks, which require fine-tuning of parameters and learning rates to achieve stringent Nash equilibrium conditions, leading to problems such as gradient explosions and mode collapse. Building upon the latest research in synthetic models DDPM (denoising diffusion probabilistic model), we propose a novel approach for 2D brain MRI image synthesis based on a lightweight denoising diffusion probabilistic model. This method improves the attention module in the denoising diffusion probabilistic model to make it more lightweight. Additionally, we adopt the smooth L1 loss function as a replacement for the traditional mean absolute error (L1 loss) by comparing the error between the 2D brain MRI images with added noise and the real noise for training the model. Finally, we validate the proposed model on the MRI Brain Tumor Classification dataset, demonstrating that it achieves high-quality synthesis results while significantly reducing the parameter count of the DDPM model.

Keywords: 2D brain MRI image; image synthesis; DDPM; light weight

ARTICLE INFO

Received: 28 July 2023
Accepted: 25 September 2023
Available online: 9 October 2023

1. Introduction

Medical images, as carriers of visual information, objectively reflect the internal structures, physiological tissue organization, and abnormalities of the human body. They are typically categorized into three modalities: CT (computed tomography) images, ultrasound images, and MRI (magnetic resonance imaging) images. In recent years, with the improvement of living standards, increased work pressure, and accelerated pace of life, the incidence of brain diseases has been rising. Brain diseases are characterized by high disability rates and high recurrence rates, which has led to a growing focus on brain disease detection and diagnosis[1]. Among the various medical imaging modalities, MRI has become a commonly used tool for examining and diagnosing brain diseases[2]. It is widely applied in routine diagnoses of brain diseases such as Parkinson's disease, epilepsy, brain tumors, and cerebral infarctions. MRI is particularly valued for its ability to provide clear tumor boundaries, high-resolution soft tissue visualization using imaging techniques like T1, T2, and DWI, as well as multi-planar anatomical structures of the brain, including gray and white matter. These characteristics make it

effective for precise tumor localization and presenting detailed organ contours, which facilitates interpretation even by non-experts[3]. Additionally, MRI imaging is not affected by issues like bone shadows or acoustic attenuation, further enhancing its diagnostic value. Consequently, MRI has gradually replaced brain CT and ultrasound images, becoming the benchmark imaging modality for diagnosing and treating brain diseases. However, in the field of MRI image processing, there may be missing or artifact-affected sequences due to equipment malfunctions, improper usage, or patient sensitivity to contrast agents. Such issues can impact the diagnosis of brain diseases using multi-sequence MRI. Furthermore, the high cost of purchasing and maintaining MRI equipment leads to a lack of data samples, further limiting in-depth research on certain brain diseases or pathological conditions. To address these challenges, using generative models for MRI image data preprocessing, augmentation, or expansion can provide accurate measurement data for medical professionals, aiding in quantitative analysis of brain diseases and improving the efficiency of medical practitioners.

In recent years, with the development of deep learning technologies such as generative adversarial networks (GANs) and variational autoencoders (VAEs), computer image virtual synthesis techniques have been applied in various fields, including clinical diagnosis, disease treatment, and medical research, attracting widespread attention from researchers. The latest high-resolution synthesized images not only provide rich visual experiences but also offer highly informative and discriminative information. A recent study proposed a GAN model composed of a generator and a discriminator[4]. The generator models the latent distribution of data and generates new samples, while the discriminator distinguishes the differences between real and generated samples to synthesize images. However, this method suffers from the problem of "mode collapse". When the generator deceives the discriminator by producing samples from a single mode, it may get stuck in a loop of generating samples that are similar to that mode, which affects the quality and diversity of the synthesized images. To overcome this issue, another study introduced Wasserstein distance to measure the differences between generated and real samples, ensuring that the discriminator satisfies certain Lipschitz continuity conditions[5]. However, this model requires relatively complex parameter adjustments and the selection of appropriate techniques like weight clipping or gradient penalization. In a successful attempt, Han et al.[6] used an improved GAN-based synthetic brain MR image generation method to successfully synthesize multi-sequence brain magnetic resonance imaging (MRI) images with low contrast, strong anatomical consistency, and variations within sequences. The research produced a set of clear 128 × 128 resolution brain MRI images. However, all the above models are based on improved GANs, which cannot completely avoid the issues of gradient vanishing and "mode collapse" encountered in traditional GANs, making it challenging to handle complex data distributions.

In order to fundamentally address the limitations of GAN models, Sohl-Dickstein et al.[7] proposed a diffusion model based on nonequilibrium thermodynamics. This model systematically and slowly adds noise to data during iterative processes, disrupting its distribution structure. Subsequently, it reconstructs the original data distribution through an inverse process to generate data. Building upon this, Ho et al.[8] further introduced the denoising diffusion probabilistic model (DDPM), which improved the mathematical computation methods of the original diffusion probabilistic model, making it applicable for high-resolution image synthesis. But traditional DDPM based on the Markovian diffusion process, requires over 1000 iterations during the reconstruction generation phase to yield high-quality generated samples, leading to excessively long inference times. Song et al.[9] proposed an efficient iterative implicit probabilistic model that employs a non-Markovian diffusion process for reverse inference, enabling faster sampling and generation of higher quality images. In the medical field, Chung and Ye[10] trained a continuous-time, time-varying fractional matching model using denoising fractional matching. During the inference stage, they used numerical SDE (stochastic differential equation) solvers and data consistency steps for iterations to achieve MRI image reconstruction. Xie and Li[11] employed a conditional DDPM defined in the measurement domain (e.g., k-space in MRI reconstruction) to perform accelerated MRI reconstruction on under-sampled MRI masks. Additionally, the study of Hu et al.[12]

proposed a Squeeze-Excitation (SE) network that enhances the information weight of critical regions through learning channel attention for each convolutional block. This results in a noticeable performance gain in deep CNN (convolutional neural network) architectures with only a minimal increase in parameters. The research of Dhariwal and Nichol[13] has shown that DDPM, as a novel image synthesis method, exhibits relatively stable training compared to generative adversarial networks (GANs) and has surpassed traditional GANs in the field of image synthesis.

However, the DDPM model is criticized for its complexity, low sampling efficiency, and long training time. In this paper, based on the above research, we propose improvements to the DDPM model for the synthesis of 2D brain MRI images. Since we do not require redundant textual information for guidance in synthesizing brain MRI images, we modify the attention module of the original DDPM model using the Squeeze-Excitation attention module[12] to achieve lightweight results. This modification not only reduces the parameter count of the model but also maintains the synthesis quality, saving GPU memory space and shortening training time. Additionally, we have improved the model's loss function and noise schedule. The model introduces noise during the iterative process controlled by time steps and establishes a mapping relationship between image feature variables and real noise by comparing the error loss between the noisy images and real noise. This improved network model not only resolves issues like gradient explosions and mode collapse commonly found in traditional GANs but also achieves lightweight performance, significantly improving the speed of data training. We conducted tests on the MRI Brain Tumor Classification dataset, and the results showed promising synthesis performance. In future research, we aim to further improve this model by synthesizing high-quality brain MRI images in different modalities to provide more comprehensive auxiliary information for clinical diagnosis, lesion localization, and surgical planning, which has significant guiding implications.

2. Methods

In this chapter, we will provide a detailed introduction to the network architecture of the lightweight denoising diffusion probabilistic model.

2.1. Denoising diffusion probabilistic model

The denoising diffusion probabilistic model is a novel image synthesis model consisting of two parameterized Markov chains. It employs variational inference techniques to generate synthesized images that match the original data distribution within a finite time. The fundamental idea is to perturb the distribution in the forward process (signal to noise) and then reconstruct the data distribution in the reverse process (noise to signal). During the forward process as shown in **Figure 1**, a given initial 2D brain MRI image distribution is perturbed $x_0 \sim q(x)$. Adding Gaussian noise β_t to each of the t time steps in the forward process until the distribution of the image approaches the prior distribution, i.e., obtaining a standard Gaussian distribution:

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

The forward process is defined as follows:

$$q(x_{1-T}|x_0) = \prod_{t=1}^{T} q(x_t|x_{t-1}) \tag{2}$$

x_t represents the addition of noisy data at time step t, $t \in \{0, 1, 2, \ldots, T\}$ and T denotes the number of times noise is added.

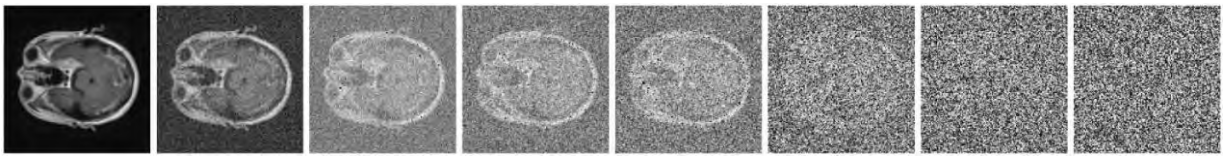

Figure 1. The forward noise addition process for the denoising diffusion probabilistic model.

In the backward iteration process, as shown in **Figure 2**, we start from a given prior and iterate using a parameterized Gaussian transition kernel. At each step t, we learn to precisely calculate the conditional probability sampling distribution $q(x_0)$, aiming to eventually recover the original data distribution.

$$p_0(x_{t-1}|x_t) = N(x_{t-1}; \mu_\theta(x_t, t), \sum_\theta(x_t, t)) \tag{3}$$

in this process, μ_θ and $\sum_\theta$ represent the mean and variance of (x_t, t).

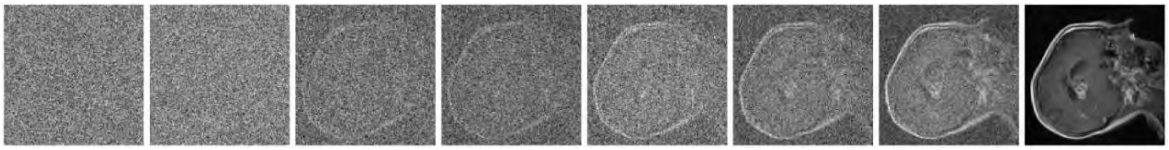

Figure 2. The backward inverse recovery process for the denoising diffusion probabilistic model.

2.2. The network structure of the lightweight denoising diffusion probabilistic model

The encoder-decoder architecture, known for its "image-to-image" transformation advantages, can effectively synthesize features by integrating implicit variables such as edges, shapes, and textures from low-level image data. Our lightweight denoising diffusion probabilistic model network structure, as shown in **Figure 3**, adopts a U-Net encoder-decoder structure[14] as the denoising and prediction model φ_z. The left part represents the encoder, which consists of four down-sampling residual blocks. Each residual block is composed of normalization layers, SiLU activation layers (compared to ReLU layers, SiLU layers are symmetric activation functions that do not saturate the negative range and can better preserve information in this region), 4 × 4 convolutional layers with a stride of 2, and attention layers. We made improvements to the original attention layer of the denoising diffusion probabilistic model by incorporating Squeeze-Excitation attention modules[12] to achieve lightweight performance, resulting in a reduction of parameters. The decoder of the network is similar to the encoder, comprising four up-sampling residual blocks. We utilize the nearest-neighbor interpolation method for upsampling in the decoder to restore the feature map's resolution to the original image size. The decoder is connected to the encoder through skip connections, where corresponding feature maps are concatenated to form richer features, enhancing the details in image synthesis.

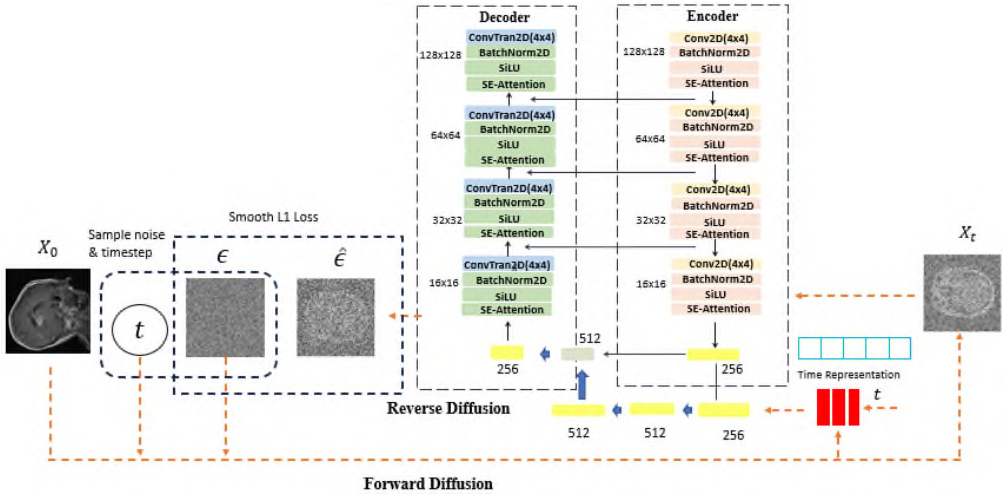

Figure 3. Lightweight denoising diffusion probabilistic model network structure.

When the input X_0 is a single-channel 128 × 128 image tensor, we utilize embedding techniques to encode the time t and noise to control the amount of noise added at each step. After passing through the encoder and decoder, the noise and image tensors are added and merged, resulting in an output tensor $\hat{\epsilon}$ of the same channel and size as the input, representing the noisy image. Finally, we calculate the loss function using the smooth L1 loss between the real noise ϵ and the noisy image tensor $\hat{\epsilon}$. By applying gradient descent, we minimize the loss function to constrain the implicit information, such as texture, shape, and style, between the noisy image feature variables and the noise. During training, the learning parameters of the entire model are updated. Eventually, we use the denoising diffusion probabilistic model reverse sampling process to recover the synthesized 2D brain MRI images.

2.3. Loss function

The DDPM model employs the L1 loss function as a measure to constrain the distance between generated images and real images. The L1 loss utilizes the mean absolute error (MAE) loss function to calculate the distance loss between the generated images and the real images, ensuring that the generated images retain their original features close to the real images during domain transformation. The expression for MAE is as follows:

$$\mathrm{MAE} = \frac{1}{MN}\sum_{p=0}^{M-1}\sum_{q=0}^{N-1}\left|z_{\mathrm{label}}(p,q) - z_{\mathrm{output}}(p,q)\right| \tag{4}$$

The gradient of the L1 loss function is constant, and in the case of a constant learning rate, the loss function will fluctuate around a stable value, making it challenging to converge to higher accuracy. Moreover, using the L1 loss in image synthesis can lead to the problem of blurring in the synthesized images' details. The smooth L1 loss, proposed in the research of Karras et al.[15], combines the advantages of both the L2 loss function and the L1 loss function, making it less sensitive to outliers. It also helps control the degree of gradient changes during training, which accelerates model convergence. Therefore, in this paper, the lightweight denoising diffusion probabilistic model adopts the smooth L1 loss function for synthesizing brain MRI images. The expression for the smooth L1 loss function is as follows:

$$\text{Smooth } L1(x) = \begin{cases} 0.5x^2 & |x| < 1 \\ |x| - 0.5 & x < -1 \text{ or } x > 1 \end{cases} \tag{5}$$

3. Experiments

3.1. Database

We validated our network model using the MRI Brain Tumor Classification dataset downloaded from the Kaggle website[16]. The MRI Brain Tumor Classification dataset is a publicly available training set for brain MRI scan image classification. It contains over 3000 brain tumor images, including pituitary adenomas, meningiomas, and gliomas, in grayscale format of transverse 2D MRI image slices as shown in **Figure 4**. The original size of these images is 256 × 256. However, due to limitations in GPU memory during experiments, we resampled each training image to a resolution of 128 × 128. The entire dataset was normalized to intensity ranges from [0, 255] to[−1, 1].

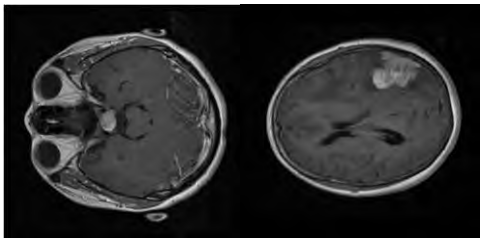

Figure 4. Scanned images from the MRI Brain Tumor Classification dataset.

3.2. Implementation details

All training and testing experiments were conducted on a server and Google Colab. The server used is Intel (R) Xeon (R) CPU e52630V4@2.20GHz. The GPU uses two NVIDIA forced ×1080 Ti with 12 GB GPU memory, with a total memory of 24 G. Google Colab provided 12 GB RAM and a GPU (graphics processing unit). The network architecture was implemented using the PyTorch framework, and the proposed model was optimized using the Adam optimizer with an initial learning rate of $lr = 5 \times 10^{-4}$. The loss function used was the smooth L1 loss as mentioned earlier. In the forward process, we adopted the cosine noise schedule strategy for adding noise. The total time steps for the diffusion model were set to 1000, and the training batch size was set to 250 epochs.

3.3. Comparisons with other method

We compared different generative network models, all of which were obtained from publicly available code. As shown in **Figure 5**, it is evident that the MRI images synthesized by GAN-based generative adversarial networks exhibit blurriness at the boundaries, along with artifacts, and poor anatomical details. The gray and white matter regions in the brain are not well-defined. In contrast, the DDPM DDIM and lightweight DDPM (LW-DDPM) models employ the image synthesis method of adding noise and denoising. This approach avoids the issues of GAN-based structures, such as failing to satisfy KL divergence Nash equilibrium and suffering from "mode collapse". The difference in image quality between DDPM and lightweight DDPM is not significant. Compared to other methods, both DDPM DDIM and lightweight DDPM are capable of sampling high-quality 2D MRI slices with clear details and realistic textures. Moreover, they exhibit higher image fidelity and outperform other generative adversarial network methods.

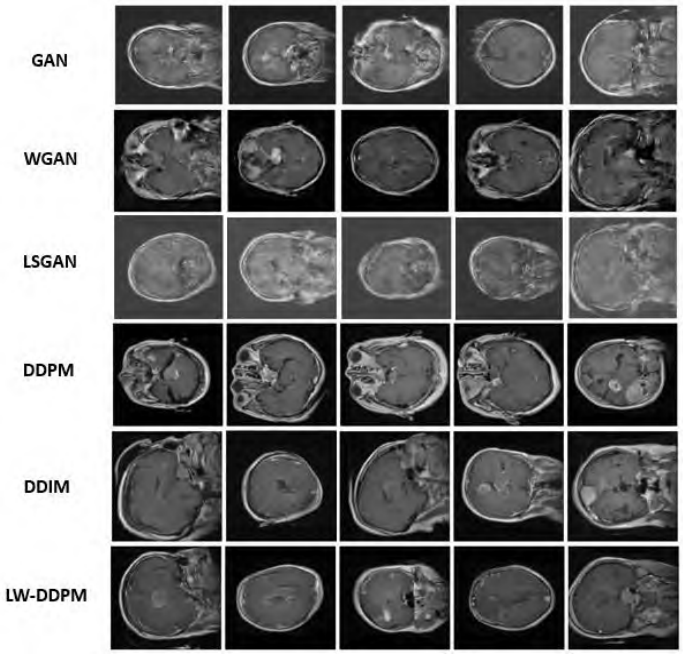

Figure 5. Comparison of the effect of different methods on synthesizing MRI images after 250 epochs.

We evaluated our synthesized images using two commonly used image synthesis metrics: structural similarity index (SSIM)[17] and Fréchet inception distance (FID). SSIM measures the similarity between two images in terms of contrast, brightness, and structural information. In this study, we used SSIM to compare the similarity between the synthesized brain MRI images and real brain images. A higher SSIM value indicates a closer resemblance between the two images. On the other hand, FID is an evaluation metric that calculates the distance between generated samples and real samples in the feature space. FID utilizes the Inception network to extract features, models the feature space with Gaussian distributions, and then calculates the distance

between the two feature spaces. A lower FID score indicates higher image quality and better diversity. **Table 1** presents the SSIM, FID, and some parameter values for different synthesis networks. The first three rows show the results for GAN-based methods, The last three rows show the results for denoising diffusion probabilistic model-based methods, From the table, it can be observed that using the denoising diffusion probabilistic model-based methods, consistently yields higher average SSIM scores and lower FIDs compared to the GAN-based methods have higher image quality and better diversity. Although the lightweight denoising diffusion probabilistic model slightly approaches DDPM in terms of FID and SSIM values, it requires fewer parameters and consumes less GPU memory, resulting in higher efficiency. The training process of DDIM is the same as that of DDPM. The reverse inference process uses a non-Markovian diffusion process with fewer steps, saving reverse inference time. However, this comes at the expense of image generation quality. To validate the statistical properties of improved model's predicted noise distribution, we employed noise distribution statistical histograms. As shown in the **Figure 6**, histograms of DDPM, DDIM, and LWDDPM predicted noise distribution and ground truth noise distribution were compared. It is evident that the statistical histogram of predicted noise distribution from LWDDPM closely aligns with the ground truth noise distribution of the image. In the **Figure 7** demonstrates that the lightweight denoising diffusion probabilistic model exhibits faster and more stable loss reduction compared to other models.

Table 1. Comparison of structural similarity (SSIM) and Fréchet inception distance (FID) results for different image synthesis methods.

Methods	**GAN**	**WGAN**	**LSGAN**	**DDPM**	**DDIM**	**LW-DDPM**
Timestep/inferencestep	/	/	/	1000/1000	1000/50	1000/1000
FID	95.65	91.27	103.86	72.24	75.43	70.02
SSIM	0.839	0.864	0.823	0.896	0.889	0.913
Parameter	/	/	/	13.48 MB	13.48 MB	13.34 MB

(/) indicates that the corresponding information was not reported in the original work or the respective parameters were not measured.

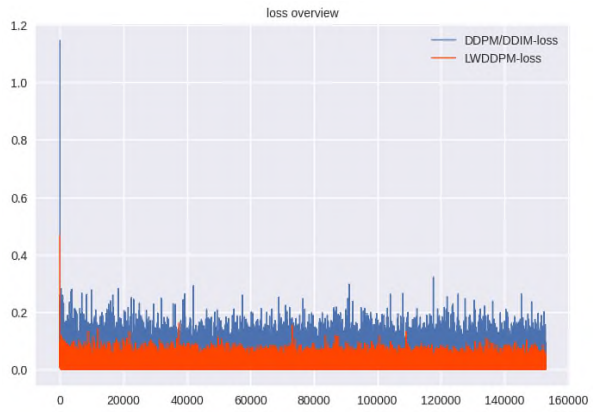

Figure 6. Comparison of statistical histograms for predicted noise distribution and ground truth noise distribution.

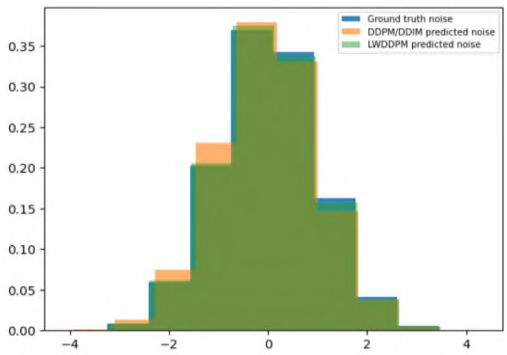

Figure 7. Loss curve for DDPM/DDIM loss and LW-DDPM loss.

We also conducted experimental comparisons on GAN, DDPM, DDIM and lightweight DDPM models for different training batch sizes: 70, 100, 200, and 250 epochs, and their synthesized results are shown in **Figure 8**. From the leftmost column, we can observe that when the training batch size is 70 epochs, the denoising model is not yet fully trained, resulting in simple distribution predictions for random noise. As a result, the MRI images restored from random noise exhibit slightly blurred texture details, and the predicted synthetic distribution is lower. The synthesized images from DDPM, DDIM and lightweight DDPM models are slightly inferior to those from the GAN-based model. At 100 epochs, although the denoising model performs well in some areas, the synthesized images still show indistinct boundaries, blurriness, and missing structural information in sensitive features. At 200 and 250 epochs, it becomes apparent that the synthesized MRI images from DDPM, DDIM and lightweight DDPM models have clearer texture details of gray and white matter anatomical structures. In contrast, the GAN-based model suffers from limitations such as gradient vanishing and mode collapse, leading to blurriness in anatomical structures and introducing a significant amount of imaging artifacts.

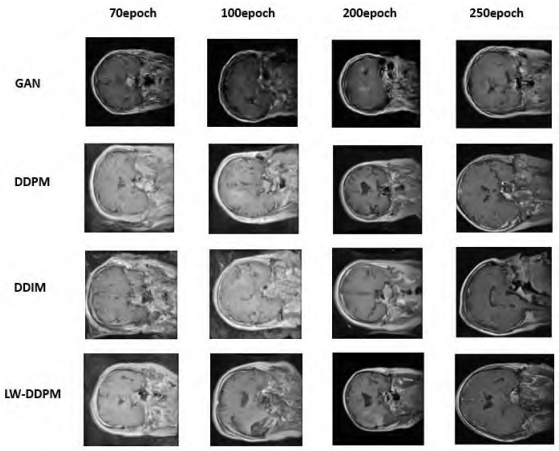

Figure 8. Comparison of synthesis effect of different training batches.

3.4. Noise schedule

Finally, we investigated the impact of different noise schedules on the synthesis results, as discussed in in the research of Nichol and Dhariwa[18]. When using the linear noise schedule in the forward process, it leads to a rapid increase in noise at lower resolutions. On the other hand, adopting the cosine noise schedule strategy in the forward process allows for better preservation of details during the noise injection intermediate steps. We tested both DDPM and lightweight DDPM models with the linear and cosine noise schedule strategies, as shown in the **Figure 9**. It was observed that when using the cosine noise schedule strategy, the synthesis performance of lightweight DDPM slightly outperformed other methods.

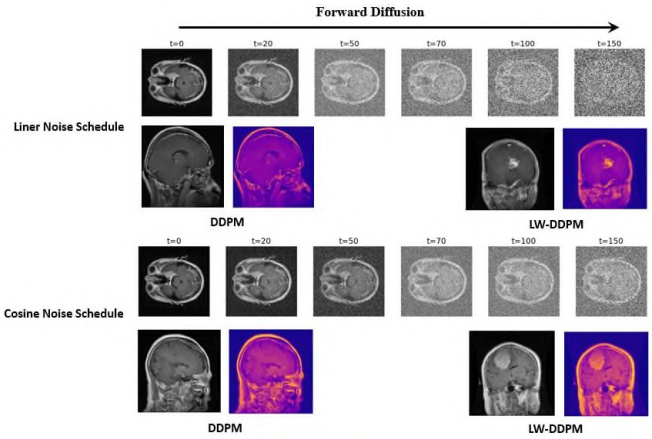

Figure 9. Comparison of synthesis effect of different noise schedule.

4. Discussion

This study applies denoising diffusion probabilistic generative models to the synthesis of 2D brain medical MRI images. The experiments compare the performance of generative adversarial networks (GANs) and denoising diffusion probabilistic models in generating 2D brain MRI images. This addresses the issue where GANs fail to improve the clarity and quality of generated brain MRI images even after multiple epoch iterations. The study also improves the original DDPM model to reduce its parameter, achieving lightweight design to save memory. This reduction in training time doesn't compromise image clarity and quality. The LW-DDPM is verified to have a predicted noise distribution that closely resembles the real image's noise distribution. Additionally, the study investigates different adding noise schedule methods and other hyperparameter adjustments.

However, this study has some limitations. The focus is placed on model enhancements and merely involves a simple division of the original MRI Brain Tumor Classification dataset into training and validation sets. Various data distribution divisions and the impact of dataset characteristics on the generated results, such as MRI imaging acquisition parameters, magnetic field inhomogeneities, and patient-specific distortions, are not considered. As a result, the generation effectiveness may not be optimal. Furthermore, the paper does not delve into how the model can control the incorporation of implicit feature information, such as image edges and textures, to generate specific features in simulated MRI images. The study also doesn't address methods for increasing sample diversity and the relationship between sample realism and diversity. In the future, these shortcomings will be addressed through further research and model enhancements.

5. Conclusion and future works

Conventional generative adversarial networks (GANs) suffer from issues such as the need for achieving KL divergence Nash equilibrium and mode collapse, resulting in poor performance in synthesizing high-resolution brain MRI images. Although the latest DDPM model can address these problems, it has higher model complexity and lower sampling efficiency, leading to longer training times. In this paper, we propose a lightweight denoising diffusion probabilistic model (LW-DDPM) for synthesizing brain MRI images. Specifically, we improved the attention module in the original DDPM model by using the Squeeze-Excitation attention module to achieve lightweight encoding. In the forward process, the encoder and decoder continuously add Gaussian noise to the dataset images. Through iterative losses, we establish constraints between the noisy data image feature variables and noise to learn the "constraint" during training. Finally, we use the trained denoising diffusion probabilistic model for reverse sampling to synthesize and recover 2D brain MRI images. Experimental results show that compared to the original DDPM, our method has fewer parameters without compromising image clarity and quality, which saves memory space and reduces training time. Furthermore, the predicted noise distribution statistics are more closely aligned with the ground truth noise distribution. In the future, we will investigate the influence of implicit feature variables in the conditional DDPM on the diffusion process and control certain implicit variables in the model to synthesize brain MRI images of different modalities with specific details. This will further enhance the realism and authenticity of synthesized medical images, providing robust assistance for clinical diagnosis by medical professionals.

Author contributions

Conceptualization, JP and GC; methodology, JP and GC; software, JP; validation, JP and GC; investigation, JP; resources, GC; data curation, JP; writing—original draft preparation, JP; writing—review and editing, JP, GC, KS and YT; visualization, JP; supervision, GC, KS and YT; project administration, GC, KS and YT; funding acquisition, GC. All authors have read and agreed to the published version of the manuscript.

Conflict of interest

The authors declare no conflict of interest.

References

1. Khan P, Kader MF, Islam SMR, et al. Machine learning and deep learning approaches for brain disease diagnosis: Principles and recent advances. *IEEE Access* 2021; 9: 37622–37655. doi: 10.1109/ACCESS.2021.3062484
2. Bakas S, Akbari H, Sotiras A, et al. Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features. *Scientific Data* 2017; 4(1): 170117. doi: 10.1038/sdata.2017.117
3. Bahadure NB, Ray AK, Thethi HP. Image analysis for MRI based brain tumor detection and feature extraction using biologically inspired BWT and SVM. *International Journal of Biomedical Imaging* 2017; 2017: 9749108. doi: 10.1155/2017/9749108
4. Goodfellow IJ, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets. *Advances in Neural Information Processing Systems* 2014; 3(11): 1–9. doi: 10.1145/3422622
5. Arjovsky M, Chintala S, Bottou L. Wasserstein generative adversarial networks. In: Proceedings of the 34th International Conference on Machine Learning; 6–11 August 2017; Sydney, NSW, Australia. pp. 214–223.
6. Han C, Hayashi H, Rundo L, et al. GAN-based synthetic brain MR image generation. In: Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018); 4–7 April 2018; Washington, DC, USA. pp. 734–738.
7. Sohl-Dickstein J, Weiss EA, Maheswaranathan N, Ganguli S. Deep unsupervised learning using nonequilibrium thermodynamics. In: Proceedings of the 32th International Conference on Machine Learning; 6–11 July 2015; Lille, France. pp. 2256–2265.
8. Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models. In: Proceedings of the 34th Conference on Neural Information Processing Systems (NeurIPS 2020); 6–12 December 2020; Vancouver, Canada. pp. 1–12.
9. Song J, Meng C, Ermon S. Denoising diffusion implicit models. In: Proceedings of the Ninth International Conference on Learning Representations (Virtual Only); 3–7 May 2021; Vienna, Austria. pp. 1–20.
10. Chung H, Ye JC. Score-based diffusion models for accelerated MRI. *Medical Image Analysis* 2022; 80: 102479. doi: 10.1016/j.media.2022.102479
11. Xie Y, Li Q. Measurement-conditioned denoising diffusion probabilistic model for under-sampled medical image reconstruction. In: Wang L, Dou Q, Fletcher PT, et al. (editors). *Medical Image Computing and Computer Assisted Intervention—MICCAI 2022*. Springer, Cham; 2022. pp. 655–664.
12. Hu J, Shen L, Sun G. Squeeze-and-excitation networks. In: Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition; 18–23 June 2018; Salt Lake City, UT, USA. pp. 7132–7141.
13. Dhariwal P, Nichol A. Diffusion models beat GANs on image synthesis. In: Proceedings of the 35th Conference on Neural Information Processing Systems (NeurIPS 2021); 6–14 December 2021; Canada. pp. 1–15.
14. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells W, et al. (editors). *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*. Springer, Cham; 2015. pp. 234–241.
15. Karras T, Aila T, Laine S, Lehtinen J. Progressive growing of GANs for improved quality, stability, and variation. In: Proceedings of the 6th International Conference on Learning Representations (ICLR 2018); 30 April–3 May 2018; Vancouver, BC, Canada. pp. 1–26.
16. Brain tumor classification (MRI). Available online: https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri (accessed on 26 September 2023).
17. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing* 2004; 13(4): 600–612. doi: 10.1109/TIP.2003.819861
18. Nichol A, Dhariwal P. Improved denoising diffusion probabilistic models. In: Proceedings of the 38th International Conference on Machine Learning; 18–24 July 2021. pp. 8162–8171.