IV. International Conference on Electrical, Computer and Energy Technologies (ICECET 2024)

25-27 July 2024, Sydney-Australia

Stain the Phony Brains: Generating Synthetic Different Medical Images Modalities Using Latent Diffusion Models

Khawla Hussein Ali

Computer Science Dept., College of Pure Science

University of Basrah

Basrah, Iraq

Khawla.ali@uobasrah.edu.iq

***Abstract*–Medical image synthesis is gaining popularity using generative models. latent diffusion models have recently shown promise in producing remarkably life-like representations of items. However, their ability to produce medical imagery has not yet been investigated. We investigate the possibility of latent diffusion models for the synthesis of different medical modalities. Initially, we create MRI images of the brains using a pre-trained DALLE-3 model given an input text prompt. Second, we create synthetic images and use 2300 MRI and 1300 CT images to train a stable diffusion model. Through a qualitative study, we assess the synthetic image data by having two impartial specialists MRI and CT modalities to classify randomly selected samples from the created data as authentic (real), fraudulent (phony), or unclear. Findings show that images produced using the diffusion model can convey qualities that are otherwise highly specific to certain medical conditions in MRI and CT modalities. This paper aims to introduce diffusion models rather than state-of-the-art generative models like Generative Adversarial Networks (GANs) and provide impetus for the medical image synthesis research community to investigate their possibilities.**

***Keywords*– Generative models, Diffusion models, Synthetic images, Artificial intelligence, MRI and CT brain mages.**

I. INTRODUCTION

Research generative models for medical image synthesis has exploded over the past ten years [1-4]. Deep generative models, such as Generative Adversarial Networks (GANs) and deep autoencoders, have demonstrated significant progress in the areas of synthesis, denoising, and super-resolution of images used in medicine [3, 5]. Numerous research has demonstrated the immense potential of GANs to produce realistic images from CT, MRI, or X-rays that can be used to train artificial intelligence (AI) models. [3,6-8]. There is growing interested in investigating the possibility of using neural diffusion models to produce medical images due to their recent success in synthesizing natural images [10, 11]. The state-of-the-art has been advanced by models like DALLE3, Mid-Journey, and Stable Diffusion for producing natural images such as art images and objects. Only the last of the three has open-source code accessible. Diffusion models are gaining popularity due to their training stability when compared to GANs. Put simply, a diffusion model is a parameterized Markov chain that has been trained by variational inference. Diffusion is used to add noise to the data so that the transition can be learned. The diffusion model, in theory, adds Gaussian noise to the input data to make it noisy, then reverses the noise to recover the data distribution. The model can produce meaningful data from random noise input once it has learned the distribution. Diffusion models, then, convert latent encoded features of picture data into more significant features. Diffusion models are comparable to denoising encoders in this sense. The entire procedure can be summed up as a two-step component, the forward process, as seen in Fig. 1. namely, the conversion of the data distribution to noise (X_i to X_T) and the reversal of the noise distribution to data distribution (X_T to X_i), or the reverse process. To train a diffusion model, one must first understand the reversing process, or $p(xt_{-1}|x_t)$. One can use the diffusion model. by performing the forward and reverse training stages on a neural network. However, the input and output dimensions of the design must match. More recent efforts have demonstrated conditioned generation by introducing guided-diffusion models, whereas before the generating capacity of diffusion models was largely exploited for the unconditional generation of data [10–12]. These projects have shown

how to create photo-realistic visuals by using the input text or image's context as a guide. Diffusion models are currently being used for text-to-image applications, or for creating visuals based on word prompts. Furthermore, Han et al. [13] introduced the Classification and Regression Diffusion Model (CARD) and illustrated its use in both regression and classification applications. In CARD, the authors used generative modeling conditioned on the class labels to tackle the problem of supervised learning. While claiming state-of-the-art outcomes was not the aim, the method's performance on the benchmark dataset has been encouraging. The model's classification accuracy for CIFAR-10 was 90.9%. One might anticipate that diffusion models will be able to produce a wide range of medical images because of their capacity to learn the representation. Additionally, they can give current methods for medical imaging applications such as domain-to-domain translation, super-resolution, noise adaptation, and noise removal, data augmentation.
There is now work being done on the synthesis of medical pictures utilizing neural diffusion models, aside from the recent paper [14]. Walter et al. [14] produced T1w MRI brain pictures using latent diffusion models. A stack of 100,000 brain MRI images conditioned on important characteristics like age, sex, and brain capacity was created by using 31,740 brain MRI and CT scans from the UK Biobank. In this study, we investigate neural networks of diffusion models.

We create these using the stable diffusion model and the DALLE3 model.
the images show them to a pair of specialists for their opinions. After that, we compile the specialists' comments and list some of the difficulties in creating medical image synthesis utilizing the neural diffusion paradigm. The remaining text is arranged as follows: Section 2 delineates our work methodology. The outcomes of creating MRI and CT brain images are shown in Section 3, as some insights into the results. Section 4 finally brings the paper to a conclusion.

II. Methodology

In this paper, we designed two experiments to create artificial brain MRI and CT modalities. In the first experiment, we generated images based on the input text by using the OpenAI DALLE3 API. The DALLE3 model has garnered significant attention lately due to its capacity to produce remarkably lifelike representations of objects based on a given input text. We created several MRI and CT images of the brains using the API. Two specialists with training were then shown a randomly chosen selection of the generated images. We requested two major duties from the specialists.
We asked them to classify each image according to their perceived understanding as real, phony, or unclear. Secondly, we requested a concise explanation of any relevant details on brain health or disease diagnosis (e.g., healthy brains with seriously damaged brains, pneumonia-affected brains, etc.). The labels of the images were unknown to the specialists beforehand. Every image we showed the specialists was artificial. The specialists worked individually and were strangers to one another. One of the two specialists was already familiar with generative models and artificial intelligence, whereas the other was unfamiliar with deep generative models.
Specialists weren't aware of applying the stable diffusion model in the second experiment [13]. Our stable diffusion model was trained with 2300 MRI images from [16] and 1300 CT images. From [17]. The images were resized to have a resolution of 256 by 256. Other pre-processing was not carried out.

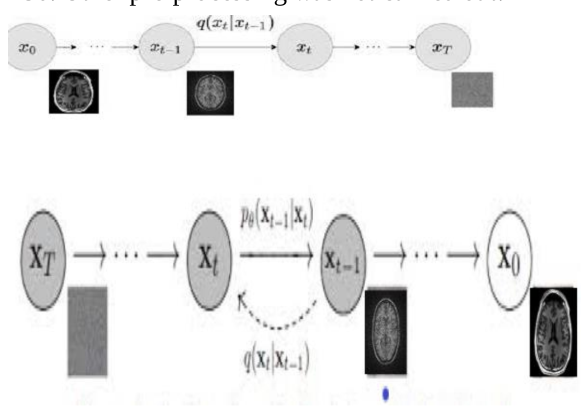

Fig. 1. Forward and reverse processes in diffusion model training. Figure modified from [9].

The forward diffusion model can be written in math. Process and it is fixed.

$$q(x_t\,|x_{t-1}) := \left(x_t;\, \sqrt{1-\beta_t}\, x_{t-1}, \beta_t I\right) \quad (1)$$

The reverse diffusion model is:

$$p\theta(x0{:}T) = p(xT)\prod_{t=1}^{T} p\theta(x_{t-1)}|x_t) \quad (2)$$

Where:

$$p\theta(x_{t-1}|x_t) = \mathcal{N}(x_{t-1}, \mu\theta(x_t, t), \sum_{0} x_t, t)) \qquad (3)$$

III. Synthetic Medical Datasets

We generated and trained 2300 MRI and 1300 CT medical images using latent diffusion models that contain two parts; forward diffusion which uses a math process for diffuse Gaussian noise in the original image, and the output is pure noise. The second part is a reversed diffusion model that takes pure noise as input and tries to remove noise and generate MRI and CT by training.

IV. RESULTS

We produced 95 images in total using the latent diffusion models. We think the sole restriction on the quantity of generated images is the tokens at our disposal. Examples of MRI and CT images are displayed in Figs. 2 and 3, in that order. Specialist A determined that 14 images were genuine out of the 40 images we showed the specialists, whereas 4 MRI and 17 CT images were false. Two MRI images were rated as unclear by Specialist A. Specialist B, the second specialist, recognized ten MRI images and two CT scans as authentic(real), whereas all other images were deemed to be phony (fraudulent). Consensus among specialists: Out of the twenty CT scans, only three were identified as authentic by both specialists. In a similar vein, both specialists certified five MRI images as authentic. It was unclear which two MRI images the specialists thought were which. In experiment 2, the specialists made some intriguing remarks when we asked them to give a quick explanation of what the images may show.
Table 1 contains a collection of some descriptions, for instance. These descriptions make it abundantly evident that the model was able to produce features that correspond to particular brain diseases, and that some of the images carried features that were identical to actual MRI images.

V. DISCUSSION

Specialists soon recognized some of the generated images as false because they lacked realistic image qualities. These images were referred to as displaying unusual exposure or having ribs that

Table I. SAMPLES OF NOTES FROM THE SPECIALIST

Image modality	Notes
MRI	this image would be considered abnormal because it shows areas that are different in appearance from what you would expect in a typical, healthy brain scan. These differences could be darker or lighter areas that might represent various issues such as a tumor, swelling, or damaged tissue.
CT	Possible damage after trauma to the head or brain tumor

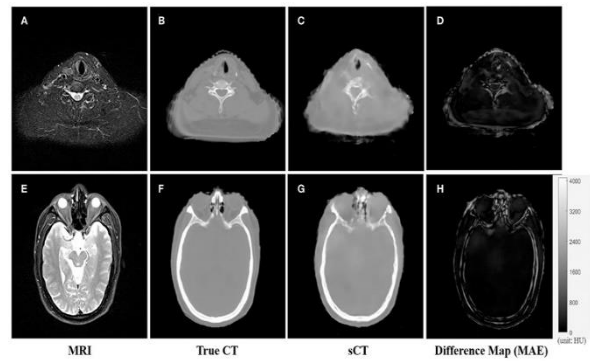

Fig. 3. Samples of brain CT images generated through the latent diffusion model

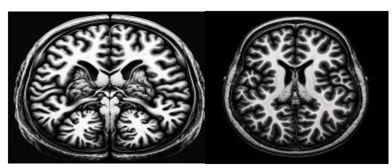

Fig.4. Samples of synthetic images for images MRI identified as phony by at least two specialists

One important finding about the phony photos was that, in contrast to actual MRI imaging, the trachea is visible behind the cardiac shadow.

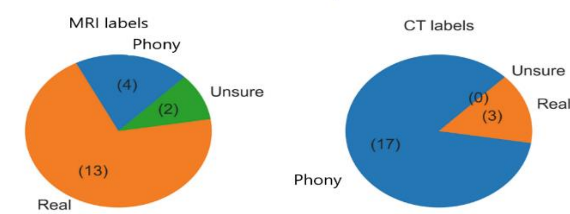

Fig. 5. Evaluation by Specialist

A. *Limitations*

Diffusion models have been found to have limitations when it comes to producing details in complicated scenes [9]. Therefore, super-resolution or noise adaption techniques would need to be used in addition to the generation of complicated medical images [5]. Diffusion model training, like many other AI models, is susceptible to bias in the dataset; for instance, imbalanced features of medical problems in the input CT or MRI image, or intrinsic noise in the data. Thus, the bias will also be present in the synthetic data derived from such a diffusion model. The bias may eventually become a cascade behavior and be further amplified if the collected data are made available to the public and used for further model training [7]. The model has not been very explainable because it has been utilized essentially as a black-box model. A few examples of pictures that at least two people said were phony as seen in Fig 4. Specialists rapidly recognized several of the generated images from the pre-trained model as false since they blatantly lacked actual image properties. These images were described as displaying unique exposure, bizarre clavicle appearance, and unusual rib look. A summary of Specialist A's evaluation can be found in Fig. 5. We have shown in this work the potential of neural diffusion models for MRI and CT image creation of the brains. Even though the specialists saw a lot of images as phony, only a small percentage of images were still classified as authentic or real. The specialists' labeling indicates that some of the produced MRI images contained striking similarities to actual photos. With the CT images, however, it was simple to identify bogus images. We demonstrated that neural diffusion models have a great deal of potential for learning intricate representations of medical pictures through qualitative inspection of the generated images. Papers on the diffusion model for medical image synthesis has not yet reached a mature state, even though latent diffusion models perform better than GANs-based techniques for creating synthetic natural images.

VI. CONCLUSION AND FUTURE WORK

We have shown in this paper that neural diffusion models have the potential to be a combination of MRI and CT medical brains. Even though the specialists saw a lot of images as fake, only a small percentage of images were still classified as authentic. The marking from the specialists indicates that certain produced MRI images contained a great resemblance to real images. With the CT images, however, it was simple to identify false images. We demonstrated that neural diffusion models have a great deal of potential for learning intricate representations of different medical images through qualitative inspection of the generated images. Research efforts on the diffusion model for medical image synthesis have not yet reached maturity, even though diffusion models perform better than GANs-based approaches for synthesizing natural images.

REFERENCES

1. Hazrat Ali, Shafaq M., Zubair Shah: Spot the Fake Lungs Generating Synthetic Medical Images using Neural diffusion models. 30th Irish Conference, AICS 2022 Munster, Ireland, December 8-9, (2022)

2. Haneen Majid, Khawla Hussein: Automatic Diagnosis of Coronavirus using Conditional Generative Adversarial Network(CGAN). Iraqi Journal of Science, (2023) Vol. 64, No. 7 pp. 3642-3656

3. Yi, X., Walia, E., Babyn, P.: Generative adversarial network in medical imaging: a review. Med. Image Anal. 58, 101552 (2019)

4. Jiang, Y., Chen, H., Loew, M., Ko, H.: Covid-19 CT image synthesis with a conditional generative adversarial network. IEEE J. Biomed. Health Inform. 25(2), 441–452 (2020)

5. Chen, M., Shi, X., Zhang, Y., Wu, D., Guizani, M.: Deep feature learning for medical image analysis with convolutional autoencoder neural network. IEEE Trans. Big Data 7(4), 750–758 (2017)

6. Ali, H., et al.: The role of generative adversarial networks in brain MRI: a scoping review. Insights Imaging 13(1), 1–15 (2022)

7. Ahmad, W., Ali, H., Shah, Z., Azmat, S.: A new generative adversarial network for medical images super-resolution. Sci. Rep. 12(1), 1–20 (2022)

8. Munawar, F., Azmat, S., Iqbal, T., Gronlund, C., Ali, H.: Segmentation of lungs in chest X-ray image using generative adversarial networks. IEEE Access 8, 153535– 153545 (2020)

9. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. In: Advances in Neural Information Processing Systems, vol. 33, pp. 6840–6851 (2020)

10. Dhariwal, P., Nichol, A.: Diffusion models beat GANs on image synthesis. In: Advances in Neural Information Processing Systems, vol. 34, pp. 8780–8794 (2021)

11. Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., Chen, M.: Hierarchical textconditional image generation with clip latent. arXiv preprint arXiv:2204.06125 (2022)

12. Song, Y., Sohl-Dickstein, J., Kingma, D.P., Kumar, A., Ermon, S., Poole, B.: Score-based generative modeling through stochastic differential equations. arXiv preprint arXiv:2011.13456 (2020)

13. Han, X., Zheng, H., Zhou, M.: CARD: classification and regression diffusion models. arXiv preprint arXiv:2206.07275 (2022)

14. Pinaya, W.H., et al.: Brain imaging generation with latent diffusion models. arXiv preprint arXiv:2209.07162 (2022)

15. O'Connor, R.: How to run stable diffusion locally to generate images. https:// www.assemblyai.com/. Accessed 01 Oct 2022

16. Chowdhury, M.E., et al.: Can AI help in screening viral and Covid-19 pneumonia? IEEE Access 8, 132665–132676 (2020)

17. Bangning Ji, Gang He, Zhengguo C., Ling Zhao; A Novel Diffusion-Model Based OCT Image Inpainting Algorithm for Wide Saturation Artifacts. Conference: Chinese on Pattern Recognition and Computer Vision (PRCV)Xiamen, China13 October 2023