

Research Article

International Journal of Digital Journalism

Structural MRI Synthesis for Alzheimer's Disease via Conditional Diffusion on Anatomical Masks

Muge Zhang¹ , Muhammad Ali Khaliq^{2*} , Jamal Alsakran³, Byeong Kil Lee⁴ and Jeeho Ryoo⁵ ^{1,3,5}Fairleigh Dickinson University, Vancouver, BC, Canada^{2,4}University of Colorado, Colorado Springs (UCCS), Colorado Springs, CO, USA***Corresponding Author**

Muhammad Ali Khaliq, University of Colorado, Colorado Springs (UCCS), Colorado Springs, CO, USA.

Submitted: 2026, May 10 **Accepted:** 2026, Jul 03; **Published:** Under Process**Citation:** Zhang, M., Khaliq, M. A., Alsakran, J., Lee, B. K., Ryoo, J. (2026). Structural MRI Synthesis for Alzheimer's Disease via Conditional Diffusion on Anatomical Masks. *Int j Digital journalism*, 2(2), 01-09.**Abstract**

Recent advances in generative machine learning models have significantly improved medical imaging, offering promising solutions for data augmentation, privacy preservation, and improved model generalization. However, synthesizing highquality structural MRI data for Alzheimer's Disease (AD) remains challenging due to the subtle, region-specific, and progressive anatomical changes associated with neurodegeneration. In this paper, we extend the Med-DDPM conditional diffusion model – originally designed for brain tumor synthesis – to generate 3D structural MRIs specifically tailored to AD. We adopted MedDDPM due to its established stability and structural fidelity compared to other generative models, which makes it particularly suitable for capturing the subtle anatomical changes characteristic of AD. Our approach conditions the diffusion process on anatomical segmentation masks derived from the ADNI dataset, incorporating key AD-relevant brain structures into the generation process. We systematically evaluate the quality and utility of the synthetic images by training segmentation models on real, synthetic, and hybrid (mixed) datasets. Experimental results demonstrate that segmentation models trained exclusively on synthetic data achieve comparable Dice scores (0.6532) to those trained on real data (0.6513), while exhibiting significantly enhanced recall. Notably, models trained on hybrid datasets (mixing real and synthetic images) outperform both real and synthetic-only baselines, achieving a Dice score of 0.7244. These findings underscore the successful use of conditional diffusion models for generating anatomically accurate, AD-specific synthetic MRIs, and highlight their potential for enhancing training data availability, improving diagnostic accuracy, and promoting research reproducibility in neuroimaging studies.

Keywords: Military ESP, Genre-Based Pedagogy, Cognitive Load Theory, Spiral Curriculum, Professional Communication Competence**1. Introduction**

Recent Machine Learning (ML) techniques have led to more accurate diagnosis, segmentation, and treatment planning in various medical domains [1]. As the maturity and popularity of ML in medical imaging have grown, obtaining diverse and high-quality training data has become increasingly important. However, in practice, medical imaging datasets – especially 3D datasets such as Magnetic Resonance Imaging (MRI) – are limited by privacy regulations, high costs of acquisition, and availability of labeled data [2]. These challenges are especially evident in neuroimaging, where detailed structural annotations are rare and expensive to obtain, and data sharing is difficult for the aforementioned reasons. Synthetic medical image generation has emerged as a strategy to address these limitations. These images are artificially generated

based on real images (not obtained from real patients), so free from various restrictions such as privacy.

Synthetic imaging methods enable the expansion of training sets and support data augmentation by producing realistic, labeled data, without raising privacy concerns [3]. Previous works have demonstrated the use of generative models in domains such as brain tumor, multiple sclerosis, and stroke, but applications in neurodegenerative diseases like Alzheimer's Disease (AD) are still not fully explored [4,5]. In particular, synthesizing realistic AD-affected brain MRIs is challenging because AD pathology is diffuse and progressive, involving subtle atrophy across multiple regions rather than large focal lesions.

In this paper, we explore the use of synthetic 3D brain MRI generation to support diagnostic and segmentation tasks relevant to AD [6]. Our method adapts the state-of-the-art structural image synthetic method, Med-DDPM, which is a conditional denoising diffusion probabilistic model originally developed for tumor imaging. We extend Med-DDPM for the AD context, focusing on generating synthetic T1-weighted brain MRIs conditioned on brain structure masks from the ADNI dataset [7].

In our implementation, we segmented anatomical brain regions that are affected in AD and used these segmentation maps as conditional inputs to train the Med DDPM model [8]. The resulting generative model produces full 3D brain MRIs with variable levels of atrophy and other disease-related features characteristic of AD. We then evaluate the quality of the generated synthetic images using a separate segmentation model trained on real data versus synthetic data versus a combination of both [4]. Our findings show that segmentation models trained on our synthetic images achieve performance comparable to or even better than those trained on real data.

In summary, we make the following contributions:

- We extended the Med-DDPM conditional diffusion model to generate 3D brain MRIs specific to AD, conditioning the generation process on segmentation masks of AD relevant brain structures.
- We demonstrate the quality of the resulting synthetic images by evaluating their impact on downstream segmentation model performance, showing that models trained on purely synthetic data can achieve similar performance to models trained on real data.
- We illustrate how synthetic MRI generation can serve as a tool for training data augmentation and privacy-conscious data sharing in neuroimaging research, particularly for diseases like AD where data are limited.

The remainder of the paper is organized as follows: Section II provides background on synthetic image generation and AD imaging; Section III details the architecture and design of our adapted synthesis pipeline; Section IV describes our experimental methodology; Section V presents experimental results and analysis; Section VI discusses recent work in the synthetic medical imaging domain; and Section VII concludes the paper with a discussion of future directions.

2. Background

Synthetic image generation is the process of creating artificially generated images using generative ML models without using a physical camera or sensor. These images are not captured from the real world, and sometimes do not look anything like the real image to human eyes, yet are very effective in training AI models and simulations. The recent emergence of generative models has facilitated the successful application of synthetic images in various domains such as medical images, autonomous driving, and creative AI [9-11]. In medical imaging, synthetic data generation has been used for tasks such as data augmentation, lesion and disease

simulation, anonymization, image denoising, and even simulating brain morphology changes due to aging [8,3].

AD poses unique challenges in neuroimaging due to its spatially diffuse pathology. Unlike tumors, which appear as discrete lesions, AD related changes such as cortical thinning and hippocampal atrophy manifest gradually across multiple regions [13]. This lack of a well-localized abnormality complicates both diagnosis and the training of ML models since the disease effects are subtle and distributed. Therefore, it is especially important and valuable to have region-level control in generative models for AD, allowing specific structures to be manipulated or emphasized. However, acquiring large, annotated datasets with detailed anatomical labels for AD is difficult due to privacy constraints, high costs, and the scarcity of expert annotations [14].

A variety of generative modeling approaches have been explored to overcome these limitations. Prominent categories include Variational Autoencoders (VAEs), Generative Adversarial Networks (GANs), and diffusion models (including latent diffusion) [15-17]. These models learn the data distribution from a set of neuroimaging scans and can then synthesize new, realistic brain images. For example, VAEs and GANs have been applied to brain MRI synthesis and augmentation tasks [18]. While these approaches have shown promise, each has notable drawbacks. VAEs can suffer from limited image fidelity and issues like posterior collapse, which lead to blurry or less varied outputs [19]. GANs, while capable of producing high-resolution images, often face training instability and mode collapse, resulting in difficulty capturing the full diversity of brain appearances. Additionally, GAN training can be sensitive and hard to reproduce, which is problematic for reliable medical image generation [20].

Diffusion-based generative models have recently emerged as a promising alternative that addresses some of these issues [21]. Diffusion models generate images by iteratively denoising random noise, which generally provides greater stability in training and the ability to capture complex data distributions. In the context of medical imaging, conditional and latent diffusion models have achieved state-of-the-art fidelity in synthesis tasks. However, straightforward applications of diffusion models in neurodegeneration research have been limited. For example, latent diffusion models trained on binary AD diagnostic labels can mimic general disease-related features but do not offer fine-grained anatomical control necessary for accurate modeling of AD pathology [7,22]. A heuristic approach such as SynthSeg can simulate images from given label maps (segmentations), but it does so without a learned generative model, meaning it may not capture realistic intensity distributions as well as a trained model [12].

In comparison to the aforementioned methods, Med-DDPM enables high-fidelity 3D MRI synthesis conditioned on structural masks [7]. Med-DDPM is a conditional denoising diffusion probabilistic model that demonstrated effective synthesis of brain tumor MRIs when provided with segmentation masks indicating tumor location

[23]. It achieved competitive segmentation performance using only synthetic training data and showed advantages over GAN-based methods in terms of training stability and output realism. However, prior implementations of Med-DDPM employed relatively simple conditioning schemes (e.g., binary tumor-versus-background masks), which are less applicable to diseases like AD where structural changes are more subtle and anatomically widespread. To our knowledge, no prior work has applied structural mask conditioned diffusion to neuro-degenerative imaging.

In this study, we extend Med-DDPM to generate AD specific brain MRIs by conditioning on detailed anatomical segmentation masks relevant to AD pathology. We use segmentation maps produced by Fast Surfer [24], a fast and fully-automated deep learning tool for brain MRI segmentation, to provide the structural conditioning. These Fast Surfer generated maps include brain regions known to be affected in AD (such as the hippocampus and ventricles) and are derived from T1-weighted MRIs in the ADNI dataset [7]. Without modifying the core Med-DDPM architecture, we

retrain the model to synthesize realistic AD-like MRIs conditioned on these anatomical masks. Our goal is to use the resulting synthetic images for training-time data augmentation, and to evaluate whether models trained on synthetic images can perform comparably to those trained on real data. We emphasize that our focus is on structural MRI, and assume that we have access to an automated segmentation method (Fast Surfer) to obtain the conditioning masks [25]. Within these bounds, this work offers a targeted, anatomy-aware adaptation of conditional diffusion for neurodegenerative disease research.

3. Design

Our system architecture is designed to determine whether synthetic structural MRI images generated using segmentation aware diffusion models can serve as a viable alternative to real data for training segmentation models in the context of AD. We build our approach on the Med-DDPM framework, adapting it for the AD application. The overall pipeline consists of three major stages (Figure 1): segmentation, synthesis, and evaluation.

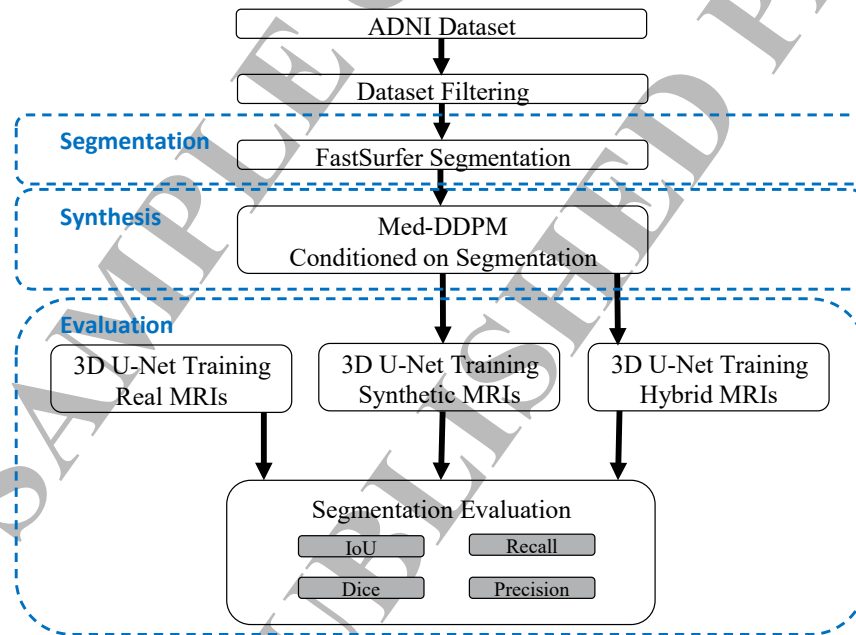


Figure 1: The Proposed Framework for Structural MRI Synthesis

In the first stage (Segmentation), we segment all T1weighted MRI images using Fast Surfer. It was chosen for its high accuracy, speed, and its use of definitions compatible with the well-established Free Surfer segmentation protocol. We ran Fast Surfer with default parameters in --seg_only mode (no model fine-tuning). This produces, for each subject, a 3D anatomical label map in the subject's native space, where each voxel is labeled as one of several brain structures. These segmentation maps serve as structural masks for training the diffusion model in the next stage. In essence, the output of Fast Surfer is a mask image where different brain regions (cortex, ventricles, hippocampi, etc.) are encoded by different label values. By using these as inputs, we

ensure the generative model is aware of the anatomical layout of each brain.

In the second stage (Synthesis), we use the segmented masks to condition Med-DDPM for image generation. We keep the original Med-DDPM architecture intact but retrain it to generate full-resolution 3D brain MRIs conditioned on the Fast Surfer segmentation maps from stage one. During training, the diffusion model learns to map anatomical structures to realistic MRI intensity patterns corresponding to AD-affected brains. This means it captures the structural variability associated with different stages of dementia. For example, AD brains typically show region-

specific atrophy, especially in the hippocampus, and enlargement of structures like the lateral ventricles due to tissue loss.

The diffusion model learns to translate a given segmentation mask into an MRI volume that reflects such changes (e.g., thinner cortex or enlarged ventricles) in a realistic way. Importantly, although the model is trained on real ADNI scans, it can generate new combinations of anatomical features that were not seen together in any single real subject, while still maintaining plausible anatomy. These novel variations can serve as additional training examples for a segmentation model, potentially helping it generalize better. Figure 2 shows representative slices from synthetic 3D T1-weighted MRIs generated by our trained Med-DDPM after 50 training epochs. The synthetic images resemble realistic brain MRIs and exhibit variable degrees of atrophy and ventricle expansion, demonstrating that the model has learned to embed AD-specific traits into the generated images.

In the final stage (Evaluation), we assess the usefulness of the synthetic data by designing three segmentation experiments. We use a standard 3D U-Net architecture for segmenting a particular target structure (or set of structures) relevant to AD (See Figure 1) [26]. The 3D U-Net is widely used in medical image segmentation because it combines low level and high-level features to accurately delineate anatomical regions, even when those structures are small or have indistinct boundaries. We train three separate U-Net models under different data regimes: one model is trained using only real ADNI MRI images (with their Fast Surfer segmentation masks as ground truth labels), another model is trained using only synthetic MRI images generated by our diffusion model (with the corresponding diffusion input masks as labels), and a third model is trained on a hybrid dataset consisting of a mix of real and synthetic images.

In all cases, the network architecture and training hyperparameters are kept consistent, so the only difference is the makeup of the

training data. This controlled experimental design allows us to isolate the impact of using generated images on segmentation performance. By explicitly conditioning on region-level segmentation during synthesis and using the same segmentation model architecture and evaluation criteria, we conduct a controlled study of the tradeoffs between data realism, data diversity, and downstream segmentation performance.

4. Experimental Setup

All experiments were conducted on a Google Cloud Virtual Machine (a a2-highgpu-1g instance) equipped with 12 vCPUs, 85 GB of memory, and a single NVIDIA A100 GPU (40 GB). This computational setup allowed us to efficiently process high-resolution 3D volumes and to train both the diffusion model and the segmentation models end-to-end within a reasonable time frame [27].

We used T1-weighted structural MRI scans from a curated subset of the ADNI dataset. Subjects were selected to ensure a balanced distribution across diagnosis groups, but genetic or societal factors were not specifically considered. We stratified subjects across three diagnostic categories: cognitively normal (CN), mild cognitive impairment (MCI), and Alzheimer's disease (AD). We further enforced that each subject contributes only a single scan to the dataset to avoid data leakage from multiple time points. After filtering and balancing, we obtained approximately 150 subjects in total to form our training set. This relatively modest dataset size was chosen to keep the training of the high-resolution 3D diffusion model tractable given our computational resources and time constraints. Note that these same subjects (each with a T1 MRI) were used in both the synthesis stage (as input to Fast Surfer and the diffusion model) and in the training of the real-data segmentation model (the hybrid model also saw these real scans alongside synthetic ones).

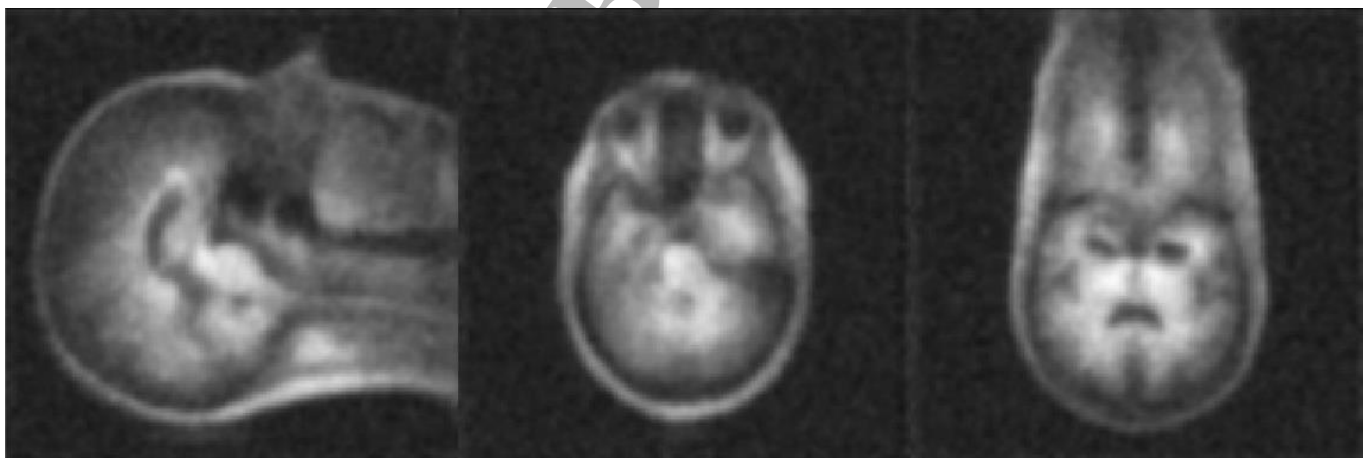


Figure 2: Synthetic T1-Weighted Brain MRI Slices Generated by Med-DDPM after 50 Training Epochs, Using Our Experimental Setup

For synthetic image generation, we employed the official Med-DDPM implementation without architectural modifications in order to preserve the model’s proven ability to generate high-quality, anatomy-conditioned 3D medical images. This decision ensures consistent performance with prior work, simplifies reproducibility, and allows us to isolate the effect of AD-specific conditioning without changing model designs. Input images and masks generated by Fast Surfer were cropped to 128×128×128 voxels and normalized [24]. This ensures compatibility with the Med-DDPM architecture, which requires input dimensions divisible by powers of two for multi-scale processing, while also capturing sufficient anatomical detail for AD-relevant regions without excessive memory overhead.

We used four conditioning channels corresponding to AD relevant anatomical regions including the hippocampus, lateral ventricles, temporal lobe, and cortical gray matter—regions strongly associated with AD progression due to their early involvement in atrophy, fluid accumulation, and structural decline [25]. These areas are consistently highlighted in clinical and neuroimaging studies as key biomarkers for AD diagnosis and monitoring. We trained the model with 155 million trainable parameters for 250 epochs. Each training epoch took approximately 18 minutes on a single NVIDIA A100 GPU, resulting in a total training time of around 75 hours for 250 epochs [26].

To assess the utility of synthetic images, we trained segmentation models using MONAI’s 3D U-Net architecture with approximately 21.3 million parameters. The network consists of four down sampling blocks and uses a hybrid Dice + Cross Entropy loss [27]. We trained three segmentation variants: (1) using real T1w MRIs and masks, (2) using only Med-DDPM generated synthetic MRIs and their corresponding masks, and (3) using a hybrid dataset (50% real, 50% synthetic). Each model was trained for 50 epochs.

5. Evaluation

To evaluate segmentation performance, we compute the Dice similarity coefficient (DSC) for AD-relevant anatomical structures, including the hippocampus, lateral ventricles, temporal lobes, and whole brain volume [29]. In addition to Dice, we also used IoU (Intersection over Union), recall, and precision as these

metrics offer complementary views of model accuracy. The Dice score measures the overlap between predicted and ground truth segmentations, capturing both false positives and false negatives and offering a balanced assessment of segmentation accuracy [8]. Therefore, it is a well used metric for medical image segmentation. IoU quantifies the ratio of the intersection to the union of predicted and actual regions, providing a stricter penalty for misalignment. Precision reflects the proportion of predicted positive pixels that are truly positive, indicating how well the model avoids over-segmentation. Recall measures the proportion of actual positive pixels correctly identified, highlighting the model’s ability to capture relevant anatomical structures without omission [30].

Performance evaluation of segmentation models is detailed in Table I. The evaluation of segmentation results shows that hybrid model performs significantly better than the real and synthetic model, while synthetic model performs slightly better than the real model. The real-only model achieved a validation Dice of 0.6513, IoU of 0.6989, precision of 0.6246, and recall of 0.6739 after 50 epochs. In comparison, the synthetic-only model achieved a Dice of 0.6532, effectively matching the real model. However, metric breakdown reveals contrasting segmentation behavior: the synthetic model yielded much higher recall (0.9743) but lower precision (0.5101), suggesting that it tends to oversegment. In a clinical setting, this behavior can be advantageous for ensuring that no AD-relevant regions are missed, which is critical for early diagnosis. However, this may also introduce false positives, potentially leading to overestimation of pathology and unnecessary follow-up procedures [31]. This likely stems from the structured, mask conditioned training dataset of Med-DDPM, which encourages completeness and spatial consistency in synthetic MRIs. This could potentially lead to cost of specificity, which means the model is more likely to include areas that do not actually belong to the target structure. While this helps ensure that important regions are not missed, it could also make the results less clear or include extra tissue that is not relevant. The contrast suggests that synthetic data appears to amplify signal around structure boundaries, increasing sensitivity to anatomical regions at risk in AD, such as the hippocampus and lateral ventricles. In practical clinical settings where false negatives are more costly than false positives, this property may be desirable.

Method	Dice	IoU	Precision	Recall
Real	0.6513	0.6989	0.6246	0.6739
Synthetic	0.6532	0.4905	0.5101	0.9743
Hybrid	0.7244	0.5810	0.6333	0.9289

Table 1: Performance Summary of Segmentation Models (Real, Synthetic, Hybrid)

The hybrid model substantially outperformed both baselines, achieving a Dice score of 0.7244, IoU of 0.5810, precision of 0.6333, and recall of 0.9289. This demonstrates that combining real and synthetic data yields complementary advantages: the biological realism of real MRIs supports generalization, while the structural diversity and label-aligned consistency of synthetic

MRIs reduce variance and promote more robust learning. Visual comparison presented in Figure 3 clearly indicates the hybrid model maintained high recall while improving precision compared to synthetic-only training, suggesting that it preserved the broader structure sensitivity of synthetic images while tempering over segmentation with real anatomical variability.

Training loss is a measure of how far the model's predictions are from the true labels during training, guiding how the model updates its weights. We used the Dice loss function, which directly optimizes for overlap between predicted and actual segmentation masks. Lower and more stable loss values during training indicate better model convergence, suggesting improved learning quality and greater potential for generalization to unseen MRI scans. Training loss curves reveal clear differences in learning dynamics across the three data regimes as shown in Figure 4. The hybrid model converges to the lowest final loss and does so more steadily and quickly than the others. This suggests that mixing real and synthetic data yields a more regularized and information-rich training signal, enabling faster generalization and more robust convergence. The synthetic-only model begins

with the highest initial loss but quickly catches up to and tracks closely with the real-only model. However, its convergence shows moderate oscillation, particularly in early epochs, reflecting less stability than the hybrid regime. This variation likely stems from differences in the intensity distribution or generative artifacts in the Med DDPM outputs, despite their strong anatomical alignment. In contrast, the real-only model exhibits more stable but slower convergence, plateauing at a higher loss—possibly due to greater anatomical noise and label variability in real data. These trends align with our segmentation results: synthetic data is sufficient to drive meaningful learning, while hybrid training amplifies the benefits of both sources, producing more stable optimization and superior performance.

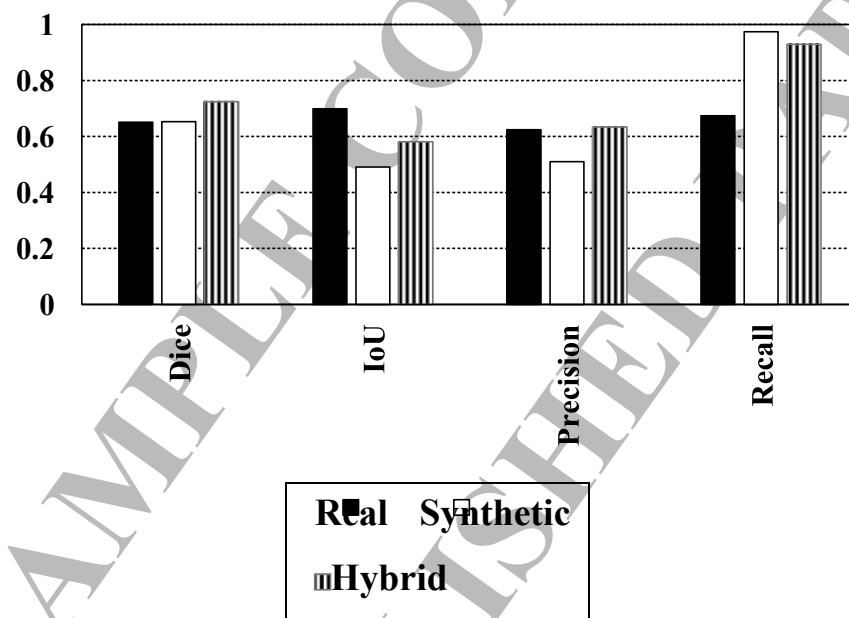


Figure 3: Comparison of Segmentation Models

Therefore, the success of the synthetic-trained and hybrid trained models validates the core hypothesis of our work: structurally-conditioned synthetic MRIs generated via diffusion models are not only anatomically plausible but also functionally effective for training segmentation models in neurodegenerative disease contexts. In other words, even though the synthetic images are computer-generated, when they are created with the proper anatomical constraints (segmentation masks), they can provide training signal of comparable quality to actual patient scans. This opens the door for greatly expanding training datasets in a privacy-preserving manner for diseases like AD, which in turn can improve the robustness and generalization of deep learning models for tasks like segmentation and possibly diagnosis.

6. Related Work

This section reviews prior work in synthetic medical image generation, focusing on three primary categories relevant to our research: Generative Models in Medical Imaging, Synthetic Data for AD, and Conditional Diffusion Models for MRI. Generative Models in Medical Imaging Early works in generative ML models utilized VAEs to capture latent representations of medical images, offering a foundation for more advanced synthetic image generation techniques [32-34]. GANs later emerged as powerful tools for synthesizing realistic medical images, particularly useful for data augmentation and anonymization [20, 35,36]. Recent techniques, including Stable Diffusion and other latent diffusion methods, have further improved image quality, stability, and scalability, enabling diverse applications ranging from lesion simulation to brain morphology modeling [16,37].

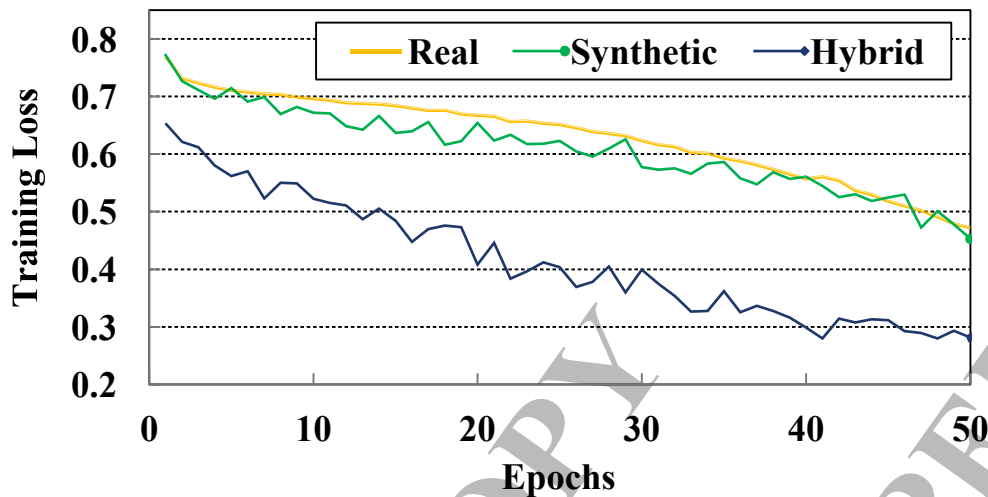


Figure 4: Training Loss Curves for Real, Synthetic and Hybrid Data

Synthetic Data for AD Applying synthetic imaging specifically to AD presents unique challenges due to the disease's diffuse and subtle anatomical changes. Previous research primarily relied on global disease labels to generate synthetic images, limiting the control of anatomical specificity [13]. Heuristic methods such as SynthSeg allowed image synthesis from label maps but lacked learned generative capabilities [12,38]. Moreover, latent diffusion models trained with binary AD labels have demonstrated the capability to capture general disease-related features but have not offered fine-grained anatomical control necessary for accurate AD modeling [6,36].

Conditional Diffusion Models for MRI Conditional diffusion models have recently emerged as promising tools for high-fidelity medical image synthesis. Notably, Med-DDPM demonstrated effectiveness in synthesizing brain tumor MRIs conditioned on structural segmentation masks, providing stability and realism superior to GAN-based methods [39]. This model has been used primarily for tumor synthesis, demonstrating improved performance for downstream segmentation tasks compared to real datasets alone [7]. However, prior implementations of Med-DDPM used simple conditioning schemes (e.g., binary segmentation masks), limiting their application to more complex and anatomically subtle diseases such as AD. Our work extends Med-DDPM by conditioning synthesis specifically on detailed anatomical segmentation masks relevant to AD pathology, enhancing anatomical specificity and improving data utility for segmentation models [20,36].

7. Conclusion and Future Work

In this paper, we presented an extension of the Med-DDPM conditional diffusion model for synthesizing high-quality 3D structural MRI images tailored specifically to AD. By conditioning the generation process on detailed anatomical segmentation masks, our method successfully produces synthetic MRIs that capture subtle, AD-specific anatomical changes. We rigorously evaluated the utility of these synthetic images in a downstream task (brain

structure segmentation) and found that models trained on the synthetic data achieved segmentation performance comparable to those trained on real data, and that a hybrid approach combining synthetic with real data yielded even better results. This demonstrates the potential of anatomically-informed diffusion models to alleviate common challenges in medical imaging related to data scarcity, privacy, and annotation costs. Future work will explore several avenues to extend this research.

- First, we plan to incorporate longitudinal data and disease progression modeling. By generating sequences of synthetic MRIs over time for the same virtual subject, we could simulate the progression of AD, which may further enhance model training for tasks like change detection or prognostic forecasting.
- Second, we aim to integrate multimodal imaging data, such as functional MRI or Positron Emission Tomography (PET) scans, into the generative process. Conditioning the diffusion model on multi-modal inputs (or generating multimodal outputs) could enrich the synthetic dataset, enabling more robust and versatile diagnostic models that leverage complementary imaging information.
- Third, we will investigate the use of our synthetic data in other downstream tasks, such as classification of AD versus MCI versus CN, to see if the benefits extend beyond segmentation. Finally, conducting extensive clinical validation of the synthetic data remains an important step. We plan to engage clinicians to qualitatively assess the realism of the synthetic MRIs and to ensure that any synthetic abnormalities correspond to plausible AD pathology. By verifying that models trained with synthetic data perform well on independent clinical datasets, we can establish greater confidence in the translational value of this approach.

In summary, our work demonstrates that conditional diffusion models, when guided by domain-specific anatomical knowledge, can generate realistic and useful medical images for a challenging application. We believe this represents a promising direction for

addressing data limitations in medical AI, and we hope it spurs further research into combining generative models with expert knowledge for enhanced data generation in healthcare.

References

1. Kumar, R. R., Shankar, S. V., Jaiswal, R., Ray, M., Budhlakoti, N., & Singh, K. N. (2025). Advances in deep learning for medical image analysis: a comprehensive investigation. *Journal of Statistical Theory and Practice*, 19(1), 9.
2. Kragsterman, P. (2024). Medical Imaging Research: 2024 Breakthroughs in AI and Advanced Technologies. *Collective Minds Radiology*.
3. Ibrahim, M., Al Khalil, Y., Amirrajab, S., Sun, C., Breeuwer, M., Pluim, J., ... & Dumontier, M. (2025). Generative AI for synthetic data across multiple medical modalities: A systematic review of recent developments and challenges. *Computers in biology and medicine*, 189, 109834.
4. Ali, M., Ali, M., Hussain, M., & Koundal, D. (2025). Generative adversarial networks (GANs) for medical image processing: recent advancements. *Archives of Computational Methods in Engineering*, 32(2), 1185-1198.
5. Czolbe, S., & Dalca, A. V. (2023). Neuralizer: General neuroimage analysis without re-training. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 6217-6230).
6. Dhinagar, N. J., Thomopoulos, S. I., & Thompson, P. M. (2024). Generative AI improves MRI-based Detection of Alzheimer's Disease by using Latent Diffusion Models and Convolutional Neural Networks. *Alzheimer's & Dementia*, 20, e089958.
7. Friedrich, P., Frisch, Y., & Cattin, P. C. (2024). Deep generative models for 3D medical image synthesis. In *Generative Machine Learning Models in Medical Image Computing* (pp. 255-278). Cham: Springer Nature Switzerland.
8. Pengfei, G., Dong, Y., Can, Z., & Daguang, X. (2024). Addressing medical imaging limitations with synthetic data generation. *Nvidia Tech. Blog*.
9. Sagers, L. W., Diao, J. A., Melas-Kyriazi, L., Groh, M., Rajpurkar, P., Adamson, A. S., ... & Manrai, A. K. (2023). Augmenting medical image classifiers with synthetic data from latent diffusion models. *arXiv preprint arXiv:2308.12453*.
10. Jian, Y., Yu, F., Singh, S., & Stamoulis, D. (2023). Stable diffusion for aerial object detection. *arXiv preprint arXiv:2311.12345*.
11. Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., & Chen, M. (2022). Hierarchical text-conditional image generation with clip latents. *arXiv preprint arXiv:2204.06125*, 1(2), 3.
12. Billot, B., Greve, D. N., Puonti, O., Thielscher, A., Van Leemput, K., Fischl, B., ... & Iglesias, J. E. (2023). SynthSeg: Segmentation of brain MRI scans of any contrast and resolution without retraining. *Medical image analysis*, 86, 102789.
13. Mulumba, J., Duan, R., Luo, B., Wu, J., Sulaiman, M., Wang, F., & Yang, Y. (2025). The role of neuroimaging in Alzheimer's disease: implications for the diagnosis, monitoring disease progression, and treatment. *Exploration of Neuroscience*, 4, 100675.
14. Nadeem, S., Azizan, P., & Michel, I. (2024). Cortex-Level Brain MRI Generation Using Diffusion Models. *Stanford University, Tech. Rep.*
15. Vivekananthan, S. (2024). Comparative analysis of generative models: Enhancing image synthesis with vaes, gans, and stable diffusion. *arXiv preprint arXiv:2408.08751*.
16. Ye, J., Ni, H., Jin, P., Huang, S. X., & Xue, Y. (2023, October). Synthetic augmentation with large-scale unconditional pre-training. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 754-764). Cham: Springer Nature Switzerland.
17. Friedrich, P., Frisch, Y., & Cattin, P. C. (2024). Deep generative models for 3D medical image synthesis. In *Generative Machine Learning Models in Medical Image Computing* (pp. 255-278). Cham: Springer Nature Switzerland.
18. Dhawan, K., & Nijhawan, S. S. (2024). Cross-modality synthetic data augmentation using gans: Enhancing brain mri and chest x-ray classification. *MedRxiv*, 2024-06.
19. Yi, X., Walia, E., & Babyn, P. (2019). Generative adversarial network in medical imaging: A review. *Medical image analysis*, 58, 101552.
20. V. Arnold. (2023). "Understanding Stable Diffusion: Advantages and Limitations.
21. Hein, D., Bozorgpour, A., Merhof, D., & Wang, G. (2025). Physics-inspired generative models in medical imaging. *Annual Review of Biomedical Engineering*, 27(1), 499-525.
22. J. Kim. (2022) "Med-DDPM: Conditional Denoising Diffusion Model for 3D Brain MRI Synthesis," IEEE Transactions on Medical Imaging, vol. 41, no. 1, pp. 123–132.
23. Henschel, L., Conjeti, S., Estrada, S., Diers, K., Fischl, B., & Reuter, M. (2020). FastSurfer-a fast and accurate deep learning based neuroimaging pipeline. *NeuroImage*, 219, 117012.
24. Czolbe, S., & Dalca, A. V. (2023). Neuralizer: General neuroimage analysis without re-training. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 6217-6230).
25. Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (pp. 234-241). Cham: Springer international publishing.
26. G.C.Platform "Google Cloud Virtual Machine Specifications," (2024).
27. Isensee, F., Jaeger, P. F., Kohl, S. A., Petersen, J., & Maier-Hein, K. H. (2021). nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature methods*, 18(2), 203-211.
28. Scherer, K. H. (2016). Experimental Setup and Methods Development. In *Grating-Based X-Ray Phase-Contrast Mammography* (pp. 37-44). Cham: Springer International Publishing.
29. Dice, L. R. (1945). Measures of the amount of ecologic association between species. *Ecology*, 26(3), 297-302.
30. Chen, S., Bai, J., Zhao, Z., Ye, T., Shi, Q., Zhou, D., ... & Qi, L. (2025). An empirical study of gpt-4o image generation

capabilities. *arXiv preprint arXiv:2504.05979*.

31. Jaccard, P. (1912). The distribution of the flora in the alpine zone. 1. *New phytologist*, 11(2), 37-50.
32. Jack Jr, C. R., Bernstein, M. A., Fox, N. C., Thompson, P., Alexander, G., Harvey, D., ... & Weiner, M. W. (2008). The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 27(4), 685-691.
33. Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 10684-10695).
34. Kingma, D. P., & Welling, M. (2013). Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*.
35. Larsen, A. B. L., Sønderby, S. K., Larochelle, H., & Winther, O. (2016, June). Autoencoding beyond pixels using a learned similarity metric. In *International conference on machine learning* (pp. 1558-1566). PMLR.
36. Rais, K., Amroune, M., Benmachiche, A., & Haouam, M. Y. (2024). Exploring variational autoencoders for medical image generation: a comprehensive study. *arXiv preprint arXiv:2411.07348*.
37. Islam, J., & Zhang, Y. (2020). GAN-based synthetic brain PET image generation: J. Islam, Y. Zhang. *Brain informatics*, 7(1), 3.
38. Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. *Advances in neural information processing systems*, 27.
39. Shin, H. C., Tenenholtz, N. A., Rogers, J. K., Schwarz, C. G., Senjem, M. L., Gunter, J. L., ... & Michalski, M. (2018, September). Medical image synthesis for data augmentation and anonymization using generative adversarial networks. In *International workshop on simulation and synthesis in medical imaging* (pp. 1-11). Cham: Springer International Publishing.

Copyright: ©2026 Muhammad Ali Khaliq, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.