Parameter-Efficient Fine-Tuning of 3D DDPM for MRI Image Generation Using Tensor Networks

Binghua Li[1,2,3,[0000−0002−2595−4762]], Ziqing Chang[2], Tong Liang[2], Chao Li[2,3], Toshihisa Tanaka[1,3], Shigeki Aoki[2], Qibin Zhao[3,1], Zhe Sun[2,⋆]

[1] Tokyo University of Agriculture and Technology, Tokyo, Japan
[2] Juntendo University, Tokyo, Japan
b.li.qr@juntendo.ac.jp, z.sun.kc@juntendo.ac.jp
[3] RIKEN Center for Advanced Intelligence Project, Tokyo, Japan

Abstract. We address the challenge of parameter-efficient fine-tuning (PEFT) for three-dimensional (3D) U-Net-based denoising diffusion probabilistic models (DDPMs) in magnetic resonance imaging (MRI) image generation. Despite its practical significance, research on parameter-efficient representations of 3D convolution operations remains limited. To bridge this gap, we propose **Ten**sor **Vo**lumetric **O**perator (TenVOO), a novel PEFT method specifically designed for fine-tuning DDPMs with 3D convolutional backbones. Leveraging tensor network modeling, TenVOO represents 3D convolution kernels with lower-dimensional tensors, effectively capturing complex spatial dependencies during fine-tuning with few parameters. We evaluate TenVOO on three downstream brain MRI datasets–*ADNI*, *PPMI*, and *BraTS2021*–by fine-tuning a DDPM pretrained on 59,830 T1-weighted brain MRI scans from the UK Biobank. Our results demonstrate that TenVOO achieves state-of-the-art performance in multi-scale structural similarity index measure (MS-SSIM), outperforming existing approaches in capturing spatial dependencies while requiring only 0.3% of the trainable parameters of the original model. Our code is available at `https://github.com/xiaovhua/tenvoo`.

Keywords: MRI image generation· Diffusion model · Parameter-efficient fine-tuning · Tensor network.

1 Introduction

Diffusion models [28], particularly denoising diffusion probabilistic models (DDPMs) [17], have gained attention for generating high-quality and diverse medical images [1,25]. In magnetic resonance imaging (MRI) tasks, these models hold great potential for enhancing clinical workflows, including disease prediction, diagnosis, and treatment in clinical practice [9,11,26]. By modeling the reverse diffusion process, DDPMs gradually remove noise to reconstruct clear images. This training paradigm offers strong robustness and controllability, enabling models to effectively capture complex data distributions. However, their large-scale architecture poses significant computational and efficiency challenges. As a consequence,

⋆ Corresponding author

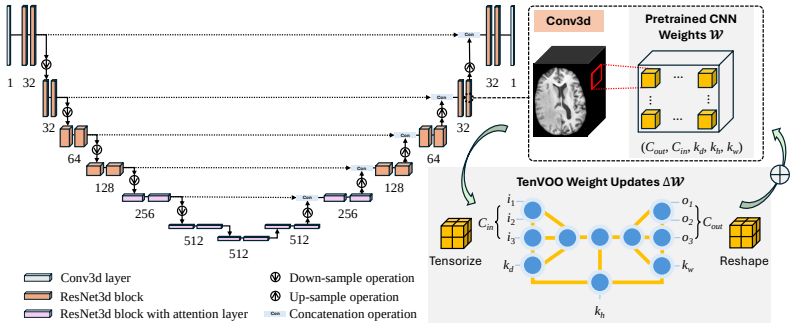

Fig. 1. Illustration of the U-Net model used for our DDPM and the TenVOO framework for 3D convolutional layers. TenVOO optimizes only the weight updates $\Delta\mathcal{W}$ via a lightweight tensor network during fine-tuning, while preserving the intricate spatial dependencies of convolution kernels throught tensor contraction.

training a DDPM typically requires high computational costs, extensive data, and resource-intensive optimization, limiting the practical applications. Meanwhile, model customization–widely desired in practice by adapting from large models–is constrained by dataset limitations, the need to store multiple parameter variations, and the complexity of designing effective concept descriptors.

To address these challenges, parameter-efficient fine-tuning (PEFT) techniques have achieved significant attention. In the existing methods, low-rank approaches [6,12,19,20] introduce trainable low-rank decompositions of model weights, serving as residual backups that can be merged during inference. Unlike adapter-based methods [18], the low-rank modeling enables efficient fine-tuning while preserving inference speed, making it particularly effective for resource-constrained inferences. However, representing conventional structures with low-rank modeling poses challenges in capturing intricate spatial relationships (see numerical results in Section 3). This limitation is particularly critical in MRI image generation, as the data typically exhibit complex spatial dependencies across multiple anatomical structures and resolution scales.

In this work, as shown in Figure 1, we propose a novel PEFT method, **Tensor Vo**lumetric **O**perator (TenVOO), for 3D U-Net [27] fine-tuning, inspired by the concept of tensor networks [3,8,21]. Compared to other low-rank counterparts, TenVOO introduces a trainable yet more compact modeling to represent complex spatial information within 3D convolutional layers, making it a powerful framework to fine-tuning U-Net-based DDPMs. Our main contributions are summarized as follows:

(1) We propose TenVOO, with two variants TenVOO-L and TenVOO-Q, for fine-tuning 3D U-Net-based DDPMs. To the best of our knowledge, TenVOO is the first application of tensor networks for fine-tuning 3D DDPM in MRI image generation.
(2) We conduct experiments on three real-world MRI datasets, including ADNI (Alzheimer's Disease Neuroimaging Initiative)[22], PPMI (Parkinson's Pro-

gression Markers Initiative)[23], and BraTS2021 [2], to fine-tune a DDPM pretrained on 59,830 T1-weighted brain MRI scans from the UK Biobank [29]. Results show that our method achieves state-of-the-art performance in MRI image generation, effectively capturing complex spatial dependencies while significantly reducing the number of trainable parameters.

2 Methodology

2.1 Preliminaries

Denoising diffusion probabilistic models (DDPMs) [17]. DDPMs are one of the most widely used diffusion models (DMs), designed to reconstruct data by modeling the reverse diffusion process within a probabilistic denoising framework. Similar to standard DMs [28], DDPMs progressively corrupt the given data $x_0 \sim p(x_0)$ with noise over T steps, following a variance schedule β_t in the forward process. The probability of the corrupted data is modeled as

$$q(x_{1:T}|x_0) = \prod_{t\geq 1} q(x_t|x_{t-1}), \qquad q(x_{x_t|x_{t-1}}) = \mathcal{N}(x_t; \sqrt{1-\beta_t}x_{t-1}, \beta_t I), \tag{1}$$

where I denotes the identity matrix. In the backword process, DDPMs iteratively remove noise from the corrupted data by estimating the mean $\mu_\theta(x_t, t)$ of the posterior as

$$p_\theta(x_{t-1}|x_t) = \mathcal{N}(x_t{-}1; \mu_\theta(x_t,t), \sigma_t^2 I), \qquad \mu_\theta(x_t,t) = \frac{1}{\sqrt{\alpha_t}}(x_t - \frac{\beta_t}{\sqrt{1-\overline{\alpha}_t}}\epsilon_\theta(x_t,t)), \tag{2}$$

where α_t represents the noise reduction factor at step t, defined as $\alpha_t = 1 - \beta_t$, and $\overline{\alpha}_t$ is the cumulative value calculated by $\overline{\alpha}_t = \prod_{s=1}^{t} \alpha_s$. Note that $\epsilon_\theta(x_t, t)$ is typically trained as a U-Net to approximate the noise in the reverse process.

Tensor networks (TNs) [10] for PEFT. TNs are a family of structural representations for high-dimensional tensors, where a tensor can be decomposed into a set of lower-dimensional core tensors interconnected through contraction operations. Following this framework, TNs can be visualized with graphs, in which core tensors correspond to nodes and index contractions are represented as edges. In the context of PEFT for example, the well-known LoRA [19] for 2D convolutional layers can be formulated as

$$\begin{gathered} \Delta\mathcal{W} = reshape(B \times_r A, [C_{out}, C_{in}, k_h, k_w]), \\ A \in \mathbb{R}^{r\times C_{in}\times k_h}, \qquad B \in \mathbb{R}^{C_{out}\times k_w \times r}, \end{gathered} \tag{3}$$

where $\Delta\mathcal{W}$ is the low-rank update of the 2D convolutional weights, r corresponds to the LoRA rank, and $C_{out}, C_{in}, k_h, k_w$ represent the number of output channels, input channels, and the kernel size in height and width, respectively. Note that the operation $\times_r$ represents the contraction between A and B along the index

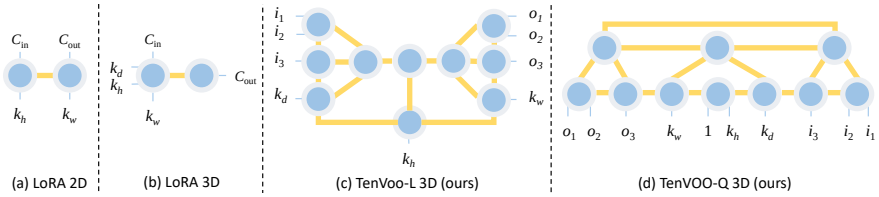

Fig. 2. Structure of (a) LoRA for 2D convolution[1]; (b) LoRA for 3D covolution[2]; (c) TenVOO-L for 3D convolution and (d) TenVOO-Q for 3D convolution. The thick yellow connections represent the tunable rank of the TN.

"r". For the $\{o, i, h, w\}$-th entry of $\Delta\mathcal{W}$, this contraction can formulated with $\Delta\mathcal{W}_{o,i,h,w} = \sum_{j=1}^{r} B_{o,w,j} A_{j,i,h}$. Furthermore, it can be visualized by a graph as shown in Figure 2 (a).

As TNs can model complex internal dependencies through contractions among low-rank tensors using few trainable parameters, they have been widely adopted in PEFT, particularly for Transformer-based architectures [21,3,8]. Among these, QuanTA [8] introduces a TN inspired by quantum circuits, facilitating high-rank representations through the contraction of a set of low-dimensional tensors, thereby preserving the representation capacity.

2.2 Problem Formulation

Recall that the objective of this work is PEFT for U-Net-based DDPMs in MRI image generation. In a standard 3D convolutional layer with ignoring the activation function, we know that the output $\mathcal{Y}$ is given by

$$\mathcal{Y} = \mathcal{W} * \mathcal{X} + b, \tag{4}$$

where $\mathcal{X}$ is the input feature map, $\mathcal{W} \in \mathbb{R}^{C_{out} \times C_{in} \times k_d \times k_h \times k_w}$ is the weight tensor, $*$ denotes the 3D convolution, and b is the bias. Since high-dimensional 3D-convolutional kernels $\mathcal{W}$ require substantial memory and computational resources, the main idea of the proposed method is thus to *efficiently represent the difference in $\mathcal{W}$ during fine-tuning using TNs*, i.e.,

$$\mathcal{Y} = (\mathcal{W} + \Delta\mathcal{W}) * \mathcal{X} + b, \qquad \Delta\mathcal{W} = reshape(\mathcal{T}_\theta(t_1, ..., t_k), [C_{out}, C_{in}, k_d, k_h, k_w]), \tag{5}$$

where $\mathcal{T}_\theta$ is the trainable TN, $\{t_1, ..., t_k\}$ represents k core tensors constituting the $\mathcal{T}_\theta$. With the updates $\Delta\mathcal{W}$ parameterized by $\mathcal{T}_\theta$, we only need to update those lower-dimensional core tensors during the fine-tuning phase.

2.3 Tensor Volumetric Operator

In this work, we propose two fine-tuning methods for 3D convolutional layers, named TenVOO-L and TenVOO-Q, inspired by LoRA [19] and QuanTA [8].

[1] LoRA implementation for 2D convolutional layer
[2] LoRA implementation for 3D convolutional layer (LoCon) from Lycoris

TenVOO-L represents the kernels using a TN structure, as illustrated in Figure 2 (c). Unlike the vanilla LoRA in Figure 2 (a), TenVOO-L first tensorizes $C_{in} = i_1 \times i_2 \times i_3$ and $C_{out} = o_1 \times o_2 \times o_3$, and then represents the weights via the TN. It is worth noting that the spatial dimensions are placed separately in TenVOO-L, enabling the model to capture spatial dependencies through tensor contraction. Moreover, when one convolutional kernel dimension is removed (e.g., k_h in Figure 2 (c)), our TenVOO-L degenerates into the 2D convolution.

TenVOO-Q, as shown in Figure 2 (d), is a direct extension of QuanTA, but with spatial dimensions explicitly assigned to different cores. Since QuanTA is renowned for maintaining high-rank representations while leveraging small-sized tensors as inputs, TenVOO-Q is expected to perform well for 3D convolutional layer with larger input or output channels, or larger kernel sizes. The corresponding number of trainable parameters for TenVOO-L and TenVOO-Q, denoted respectively $\#P_L$ and $\#P_Q$ is given as

$$\begin{aligned} \#P_L &= (i_1 i_2 + o_1 o_2) r^2 + (i_3 + o_3 + k_d + k_h + k_w + 1) r^3 + 2r^4, \\ \#P_Q &= (i_1 i_2 + o_1 o_2 + k_h) r^2 + (i_3 + o_3 + k_d + k_w) r^3 + 3r^4. \end{aligned} \tag{6}$$

We can see from Eqs. (6) that the parameter scale using TenVOO is primarily determined by r, which corresponds to the TN ranks, controling the trade-off between representation power and parameter efficiency in the model.

Linear layers follows QuanTA. Since our DDPM also includes linear layers, our TenVOO directly integrates QuanTA [8] for fine-tuning. Following their approach, we tensorize both the input and output feature dimensions into third-order tensors and then construct the QuanTA network to model these layers. Note that the linear layers account for only about 10% of the trainable parameters in our DDPM.

2.4 Initialization Is Important for TenVOO

To ensure stability in training, PEFT typically requires that the weight updates in the adapted model remain at zero during initialization. Although the core tensors in TenVOO can follow the same routine, we empirically found that this routine negatively impacts the performance of fine-tuning in practice. To address this, we adopt the initialization method of QuanTA for the convolutional layers, introducing a frozen copy of the trainable TN for initialization. In particular, given a trainable TN $\mathcal{T}_\theta(t_1, ..., t_k)$, we define its frozen copy $\mathcal{T}^{❄}$ at the beginning of training as:

$$\mathcal{T}^{❄} = \mathcal{T}_\theta(t_1^{(0)}, ..., t_k^{(0)}), \tag{7}$$

where $\{t_1^{(0)}, ..., t_k^{(0)}\}$ are the initialized core tensors before training. The weight updates during fine-tuning is then modified as

$$\begin{aligned} \mathcal{Y} &= (\mathcal{W} - \mathcal{T}^{❄} + \Delta\mathcal{W}) * \mathcal{X} + b, \\ \Delta\mathcal{W} &= reshape(\mathcal{T}_\theta(t_1, ..., t_k), [C_{out}, C_{in}, k_d, k_h, k_w]). \end{aligned} \tag{8}$$

Here only core tensors $\{t_1^{(0)}, ..., t_k^{(0)}\}$ in $\Delta\mathcal{W}$ are updated during fine-tuning. For linear layers, they follow similar initialization methods.

3 Experiments and Results

3.1 Dataset and Experimental Settings

We use 59,830 T1-weighted MRI scans from UK Biobank (UKB) [29] to pre-train an unconditional U-Net-based DDPM structured as in Figure 1, and 979, 545 and 327 T1-weighted scans from PPMI [23], ADNI [22] and BraTS2021 [2] for fine-tuning, respectively. For **UKB**, all scans are officially resampled to 1 mm^3, skull-stripped, registered to the MN152 space. For **BraTS2021**, scans are also officially resampled to 1 mm^3 and skull stripped. We follow existing work [11] to filter 327 images out of 1,251 training samples based on visual quality. For both **PPMI** and **ADNI**, the scans are resampled to 1 mm^3, skull stripped, bias corrected, tissue-segmented, registered to the MN152 space, and normalized, with all preprocessing steps performed using the CAT12 [13] toolbox.

For pre-training, we build our DDPM using MONAI [5] and train it on 59,830 UKB images. For fine-tuning, we adapt only the ResNet convolution layers, attention value/query projections, and the time embedding/projection layers for PEFT methods. We use 90% of samples for training and 10% for evaluation. Images from UKB, PPMI, and ADNI are padded and cropped to $160 \times 224 \times 160$, and BraTS2021 to $192 \times 192 \times 144$ then all resized to $128 \times 128 \times 96$. We train our DDPM using Mean Squared Error (MSE) loss, with a learning rate of 10^{-5}, batch size 1 with 4 gradient accumulation steps, and the Adam optimizer. We set the tensor rank to 4 and design TenVOO to have the fewest trainable parameters among all the baselines. For evaluation, we assess generation quality using Fréchet Inception Distance (FID) [16] and Maximum Mean Discrepancy (MMD) [15], employing Med3D [7] as the encoder following prior work [11,26]. Spatial similarity is measured using Multi-Scale Structural Similarity Index Measure (MS-SSIM) [30]. We also implement convolutional PEFT baselines, including LoRA [19], LoKr [12], and LoHa [20], following Lycoris [31].

3.2 Results

Table 1 presents the results of both generation quality and structural similarity on the evaluation set. Our TenVOOs achieve competitive performance in terms of FID and MMD compared to the baselines while outperforming them on MS-SSIM. Notably, we observe that TenVOOs yield much lower FID and MMD scores on PPMI. We speculate that T1-weighted MRI scans for Parkinson's disease exhibit minimal structural variations [4], making it challenging to adapt only the TNs in our method, especially those used for convolutions, which are designed to effectively capture spatial information, to represent them well. Additionally, TenVOOs significantly improve performance on BraTS2021, which differs the most from the pre-training data in the UKB dataset.

For a more comprehensive comparison, we also present results from jointly fine-tuned models, as given in Table 2. In this setting, in addition to adapting the target modules with low-rank representations, other layers, such as up-sampling and down-sampling layers, skip connections, and others, are also updated using

Table 1. Evaluation of generation quality using FID↓ and MMD↓, along with structural similarity measured by MS-SSIM↑. Here, Full-FT means full fine-tuning, MS-SSIM is denoted as "MS", and "#P" represents the number of trainable parameters (in millions). All FID and MMD values are scaled by 10^{-2}. The best values for each metric are highlighted in **bold**.

Model	#P (M)	ADNI			PPMI			BraTS2021		
		FID	MMD	MS	FID	MMD	MS	FID	MMD	MS
Real	–	0.016	–	0.928	0.014	–	0.934	0.004	–	0.726
Full-FT	166.67	12.001	7.760	0.701	19.751	16.177	0.646	4.006	1.366	0.436
LoRA [19]	0.82	16.794	12.729	0.511	**14.609**	**10.482**	0.431	8.648	4.330	0.076
LoKr [12]	0.79	14.626	10.460	0.434	15.857	11.881	0.253	3.399	1.086	0.291
LoHa [20]	0.82	17.031	13.166	0.490	14.677	10.529	0.426	8.759	4.452	0.084
TenVOO-L	0.60	17.349	13.331	**0.663**	16.605	12.524	0.668	**3.168**	**0.908**	**0.581**
TenVOO-Q	0.58	**13.475**	**9.502**	0.504	19.585	15.700	**0.756**	3.234	0.934	0.575

Table 2. Evaluation results for jointly fine-tuning. All settings of the table are the same as those in Table 1.

Model	#P (M)	ADNI			PPMI			BraTS2021		
		FID	MMD	MS	FID	MMD	MS	FID	MMD	MS
LoRA [19]	47.59	13.856	9.471	0.743	21.950	18.833	0.607	4.444	1.499	0.519
LoKr [12]	47.55	**11.594**	**7.359**	0.211	20.472	17.081	0.494	5.355	2.087	0.211
LoHa [20]	47.59	12.502	8.257	0.699	23.032	20.171	0.613	4.740	1.609	0.436
TenVOO-L	47.36	12.153	7.899	**0.804**	11.862	**7.539**	0.771	**1.256**	**0.190**	**0.715**
TenVOO-Q	47.34	12.676	8.756	0.507	**11.857**	7.594	**0.837**	2.594	0.737	0.462

the standard full fine-tuning approach. The overall visual results are illustrated in Figure 3. We can see that images generated by jointly fine-tuned models retain the visual structure, while those from the baseline models exhibit varying degrees of structural distortion. In contrast, our TenVOOs effectively preserve the spatial integrity of brain MRI scans.

3.3 Ablation Study

Rank. We conduct an ablation study on the tensor rank of our TenVOO. Specifically, we vary the rank across 1,2,4,6 and evaluate its impact on #P and MS-SSIM. Figure 4 (a) shows that our TenVOO can maintain a relatively low number of trainable parameters as the rank increases. Figure 4 (b) demonstrates that our TenVOO effectively learns representations for spatial information, as indicated by the increasing MS-SSIM scores with higher ranks.

Table 3. Results under different initialization strategies. "Zero" refers to zero initialization, and "Frozen" refers to the frozen-copy method.

Dataset	Method	FID ↓	MMD ↓	MS ↑
ADNI	Zero	**16.582**	**12.438**	0.543
	Frozen	17.349	13.331	**0.663**
PPMI	Zero	**14.812**	**10.767**	0.503
	Frozen	16.605	12.524	**0.668**
BraTS2021	Zero	**2.835**	**0.803**	0.409
	Frozen	3.168	0.908	**0.581**

Initialization. We compare frozen-copy and zero initialization using TenVOO-L, where the last core tensor is set to zero to disable the TenVOO backup at the start. As shown in Table 3, the frozen-copy strategy achieves better performance

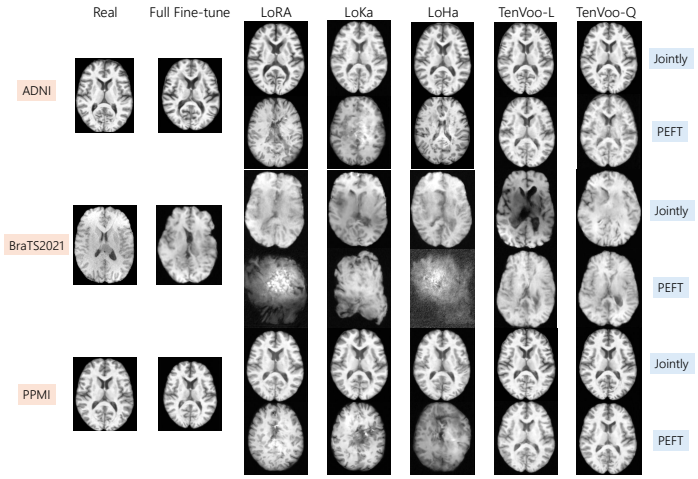

Fig. 3. Visualization results. For each image pair, the top scan is generated by the jointly fine-tuned model, while the bottom scan is generated by the adapted model.

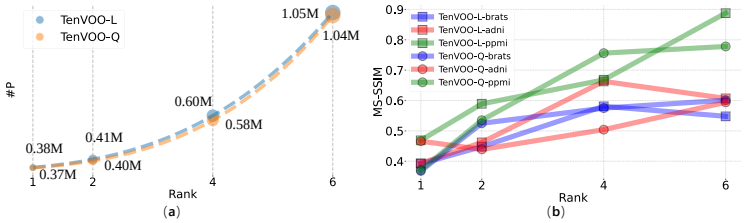

Fig. 4. Results of the ablation study on rank. (a) illustrates how the number of parameters changes as the rank increases. (b) demonstrates the impact of rank on MS-SSIM performance.

to capture structural information by providing more reasonable initialization and preserving structural information, thereby ensuring greater stability during early fine-tuning.

3.4 Efficiency Analysis

To evaluate TenVOO's efficiency, we compare it with other PEFT baselines in terms of inference time, training time, and peak allocated memory in Table 4. While TenVOO reduces parameters and maintains inference speed, memory savings are limited, likely because activation maps dominate memory usage in convolutional layers. This aligns with prior work, where memory reduction often relies on compressing activation maps [14,24]. This is further validated by results on the BraTS2021 dataset, as smaller inputs lead to smaller activation maps and greater memory savings. Combining PEFT with activation map compression may offer a more effective memory-saving strategy.

Table 4. Comparison of memory and training efficiency across PEFT methods. All values are reported as percentages relative to full fine-tuning (Full-FT).

Model	#P (%)	**UKB/ADNI/PPMI** 160 × 224 × 160			**BraTS2021** 128 × 128 × 96		
		Inference	Training	Memory	Inference	Training	Memory
Full-FT	100%	100%	100%	100%	100%	100%	100%
LoRA [19]	0.49%	100%	102.23%	94.89%	100%	117.24%	84.58%
LoKr [12]	0.47%	100%	104.55%	94.89%	100%	125.86%	84.58%
LoHa [20]	0.49%	100%	105.19%	95.04%	100%	125.86%	84.57%
TenVOO-L	0.36%	100%	118.18%	95.05%	100%	172.41%	84.82%
TenVOO-Q	0.35%	100%	120.45%	95.24%	100%	179.31%	85.38%

4 Conclusion

We introduced TenVOO, a novel parameter-efficient fine-tuning (PEFT) method for adapting 3D convolutional layers in U-Net-based DDPMs for MRI image generation. By leveraging tensor networks, TenVOO efficiently captures complex spatial dependencies while significantly reducing trainable parameters. Experimental results across multiple datasets demonstrate its effectiveness in improving structural similarity and generation quality, making it a promising approach for optimizing 3D convolutional models in medical imaging. Future work includes extending TenVOO to broader generative tasks and further refining its efficiency for large-scale applications.

Acknowledgments. This research was conducted using the UK Biobank Resource under Application Number 135200. It was also supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI under Grant-in-Aid for Scientific Research B (23K28109, 24K03005) and Grant-in-Aid for JSPS Fellows (25KJ1215).

Disclosure of Interests. The authors declare no competing interests relevant to the content of this article.

References

1. Ali, H., Murad, S., Shah, Z.: Spot the fake lungs: Generating synthetic medical images using neural diffusion models. In: Irish Conference on Artificial Intelligence and Cognitive Science (AICS). pp. 32–39 (2022)
2. Baid, U., Ghodasara, S., Mohan, S., Bilello, M., Calabrese, E., Colak, E., Farahani, K., Kalpathy-Cramer, J., et al.: The RSNA-ASNR-MICCAI BraTS 2021 benchmark on brain tumor segmentation and radiogenomic classification. arXiv preprint arXiv:2107.02314 (2021)
3. Bershatsky, D., Cherniuk, D., Daulbaev, T., Mikhalev, A., Oseledets, I.: LoTR: Low tensor rank weight adaptation. arXiv preprint arXiv:2402.01376 (2024)
4. Burton, E.J., McKeith, I.G., Burn, D.J., Williams, E.D., O'Brien, J.T.: Cerebral atrophy in Parkinson's disease with and without dementia: a comparison with Alzheimer's disease, dementia with lewy bodies and controls. Brain **127**(4), 791–800 (2004)

5. Cardoso, M.J., Li, W., Brown, R., Ma, N., Kerfoot, E., Wang, Y., Murrey, B., Myronenko, A., et al.: MONAI: An open-source framework for deep learning in healthcare. arXiv preprint arXiv:2211.02701 (2022)
6. Chavan, A., Liu, Z., Gupta, D., Xing, E., Shen, Z.: One-for-all: Generalized LoRA for parameter-efficient fine-tuning. arXiv preprint arXiv:2306.07967 (2023)
7. Chen, S., Ma, K., Zheng, Y.: Med3D: Transfer learning for 3D medical image analysis. arXiv preprint arXiv:1904.00625 (2019)
8. Chen, Z., Dangovski, R., Loh, C., Dugan, O., Luo, D., Soljačić, M.: QuanTA: Efficient high-rank fine-tuning of LLMs with quantum-informed tensor adaptation. In: Neural Information Processing Systems (NeurIPS) (2024)
9. Chung, H., Lee, E.S., Ye, J.C.: MR image denoising and super-resolution using regularized reverse diffusion. IEEE Transactions on Medical Imaging **42**(4), 922–934 (2022)
10. Cichocki, A.: Tensor networks for big data analytics and large-scale optimization problems. arXiv preprint arXiv:1407.3124 (2014)
11. Dorjsembe, Z., Pao, H., Odonchimed, S., Xiao, F.: Conditional diffusion models for semantic 3D brain MRI synthesis. IEEE Journal of Biomedical and Health Informatics (2024)
12. Edalati, A., Tahaei, M., Kobyzev, I., Nia, V.P., Clark, J.J., Rezagholizadeh, M.: Krona: Parameter efficient tuning with Kronecker adapter. arXiv preprint arXiv:2212.10650 (2022)
13. Gaser, C., Dahnke, R., Thompson, P.M., Kurth, F., Luders, E., the Alzheimer's Disease Neuroimaging Initiative: CAT: a computational anatomy toolbox for the analysis of structural MRI data. Gigascience **13**, giae049 (2024)
14. Georgiadis, G.: Accelerating convolutional neural networks via activation map compression. In: Conference on Computer Vision and Pattern Recognition (CVPR). pp. 7085–7095 (2019)
15. Gretton, A., Borgwardt, K.M., Rasch, M.J., Schölkopf, B., Smolaz, A.: A kernel two-sample test. The Journal of Machine Learning Research **13**(1), 723–773 (2012)
16. Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., Hochreiter, S.: GANs trained by a two time-scale update rule converge to a local Nash equilibrium. In: Neural Information Processing Systems (NeurIPS) (2017)
17. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. In: Advances in neural information processing systems (NeurIPS). pp. 6840–6851 (2020)
18. Houlsby, N., Giurgiu, A., Jastrzębski, S., Morrone, B., de Laroussilhe, Q., Gesmundo, A., Attariyan, M., Gelly, S.: Parameter-efficient transfer learning for NLP. In: International Conference on Machine Learning (ICML). pp. 2790–2799 (2019)
19. Hu, E., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Chen, W.: LoRA: Low-rank adaptation of large language models. In: International Conference on Learning Representations (ICLR) (2022)
20. Hyeon-Woo, N., Ye-Bin, M., Oh, T.: Fedpara: Low-rank hadamard product for communication-efficient federated learning. International Conference on Learning Representations (ICLR) (2022)
21. Jie, S., Deng, Z.: Fact: Factor-tuning for lightweight adaptation on vision transformer. In: Association for the Advancement of Artificial Intelligence (AAAI). pp. 1060–1068 (2023)
22. Jr, C.R.J., Bernstein, M.A., Fox, N.C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P.J., et al.: The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. Journal of Magnetic Resonance Imaging **27**(4), 685–691 (2008)

23. Marek, K., Jennings, D., Lasch, S., Siderowf, A., Tanner, C., Simuni, T., Coffey, C., Kieburtz, K., et al.: The Parkinson progression marker initiative (PPMI). Progress in Neurobiology **95**(4), 629–635 (2011)
24. Nguyen, L., Quélennec, A., Tartaglione, E., Tardieu, S., Nguyen, V.: Activation map compression through tensor decomposition for deep learning. In: Advances in neural information processing systems (NeurIPS) (2024)
25. Packhäuser, K., Folle, L., Thamm, F., Maier, A.: Generation of anonymous chest radiographs using latent diffusion models for training thoracic abnormality classification systems. In: International Symposium on Biomedical Imaging (ISBI). pp. 1–5 (2023)
26. Pinaya, W.H., Tudosiu, P., Dafflon, J., Costa, P.F.D., Fernandez, V., Nachev, P., Ourselin, S., Cardoso, M.J.: Brain imaging generation with latent diffusion models. In: MICCAI Workshop on Deep Generative Models. pp. 117–126 (2022)
27. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI). pp. 234–241 (2015)
28. Sohl-Dickstein, J., Weiss, E.A., Maheswaranathan, N., Ganguli, S.: Deep unsupervised learning using nonequilibrium thermodynamics. In: International Conference on Machine Learning (ICML). pp. 2256–2265 (2015)
29. Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., et al.: UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Medicine **12**(3), e1001779 (2015)
30. Wang, Z., Simoncelli, E.P., Bovik, A.C.: Multiscale structural similarity for image quality assessment. In: Asilomar Conference on Signals, Systems & Computers. vol. 2, pp. 1398–1402 (2003)
31. Yeh, S., Hsieh, Y., Gao, Z., Yang, B.B., Oh, G., Gong, Y.: Navigating text-to-image customization: From lyCORIS fine-tuning to model evaluation. In: International Conference on Learning Representations (ICLR) (2024)