PDMambaNet: Poisson Denoising-Aided Twin-Path Mamba for Brain MRI Image Segmentation*

Dayong Ren[1,2], Feifei Zhang[3,4], Fei Shi[3,4], and Aoxue Chen[1,*]
[1] Department of Neurology, Jiangsu Province Hospital,
The First Affiliated Hospital of Nanjing Medical University, Nanjing 210029, China.
[2] National Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210023, China.
[3] Xinjiang University, School of Computer Science and Technology, Xinjiang Urumqi 830046, China.
[4] Xinjiang University, Key Laboratory of Signal Detection and Processing, Xinjiang Urumqi 830046, China.

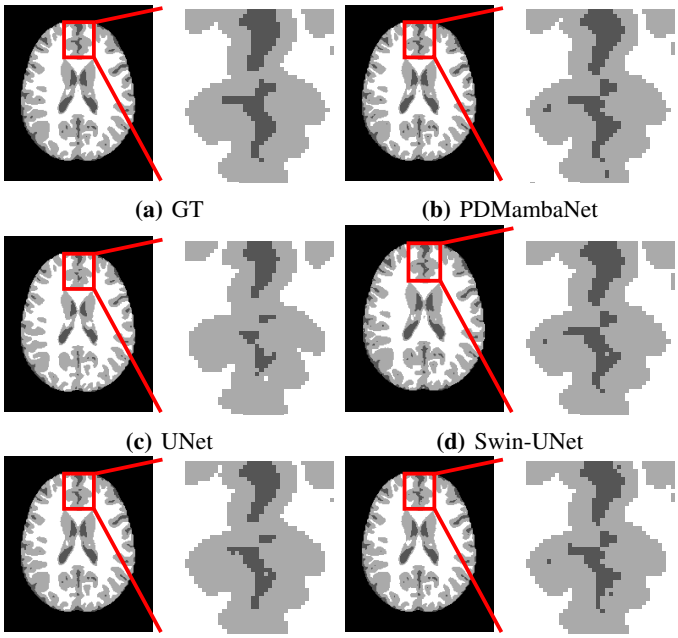
***Abstract*—Brain tissue segmentation is critical for diagnosing and treating brain diseases, but noise introduced during MRI image acquisition can compromise downstream tasks. To address this, we introduce a Poisson denoising module to eliminate Poisson and mixed noise. However, Poisson denoising may blur local brain tissue details and edges, impacting segmentation accuracy. To overcome this, we propose PDMambaNet, which integrates Poisson denoising, the Mamba architecture, and a Twin-Path Decoder (TPD). One decoder focuses on global detail recovery, while the other restores texture and edge information. This structure minimizes detail loss and improves the model's ability to capture multi-level features. The collaborative effect of TPD reduces over-smoothing and artifact generation, ensuring the preservation of cross-scale spatial information. Extensive subjective and objective evaluations demonstrate that PDMambaNet outperforms existing methods in segmentation performance. The code is available in the supplementary material.**

***Index Terms*—Segmentation, Possion denoising, Mamba, TPD**

I. Introduction

Medical images often contain Poisson noise, Gaussian noise, and impulse noise [1]. These noises can lead to false and missed detections in downstream segmentation and detection tasks, resulting in lower model accuracy. Therefore, denoising medical images before performing downstream segmentation tasks can ensure the accuracy of segmentation and detection. Existing medical image denoising methods typically model noise as zero-mean additive Gaussian noise [2], [3]. However, this assumption is not applicable in medical imaging systems, because sensor noise is usually proportional to signal strength and is more suitable for modeling with Poisson processes [4]–[7]. Removing noise is a key step in improving the performance of downstream segmentation models. These segmentation models not only improve diagnostic accuracy but also provide reliable decision support for medical professionals [8], [9].

Existing methods are based on convolutional neural networks (CNNs) and Vision Transformers (ViTs) [10]–[23]. CNN-based methods effectively capture local features but cannot efficiently utilize global contextual information, leading to suboptimal segmentation accuracy. ViTs, with their outstanding global context extraction capabilities, are widely used in medical image segmentation. However, their quadratic computational complexity results in high costs when dealing with high-resolution medical images. In resource-constrained medical environments, this high computational cost makes model deployment challenging. Recently, the development of state space models (SSMs) [24] has provided new approaches for medical image segmentation algorithms. Due to SSMs' linear complexity and excellent long-range relationship modelling capabilities, Mamba's models are widely used in medical image segmentation [25]–[31]. Mamba-based models can reduce complexity while maintaining superior segmentation



(a) GT **(b)** PDMambaNet **(c)** UNet **(d)** Swin-UNet **(e)** PNet **(f)** Mamba-UNet

Fig. 1: Segmentation results of PDMambaNet and other models on the OASIS1 dataset with zoomed-in views. As shown in the zoomed-in views, the edge structures segmented by PDMambaNet are more precise, retaining more detailed information while exhibiting smoother edge transitions.

* Corresponding author. This work was supported by a grant from the Sci-Tech Innovation Foundation for Excellent Overseas Returnees of Nanjing, China; a grant from the Outstanding Postdoctoral Foundation of Jiangsu Province, China (2023ZB500); a grant from Young Scholars Fostering Fund of the First Affiliated Hospital of Nanjing Medical University.

performance, making them widely applicable in medical image tasks [32].

To effectively remove noise from MRI images and reduce the information loss caused by denoising, this paper designs a Mamba network based on Poisson denoising and a dual-branch decoder (PDMambaNet). First, a Poisson denoising module is introduced to enhance the performance of downstream segmentation tasks. However, Poisson denoising can lead to the loss of fine details and edge blurring in MRI images. This can prevent segmentation networks from fully utilizing multi-level information, mainly when dealing with brain MRI images with complex textures and structures, potentially resulting in missed and incorrect detections. Additionally, excessive smoothing during the denoising process may weaken the segmentation network's ability to differentiate between various tissue regions, affecting the final segmentation results. To effectively mitigate the impact of denoising on downstream segmentation tasks, this paper proposes a segmentation model, MambaNet, featuring a Twin-Path Decoder (TPD) branch structure. In MambaNet, one decoder focuses on recovering global fine details, while the other is dedicated to extracting and preserving texture and edge information. This structure helps reduce the loss of details and enhances the model's ability to capture multi-level information. The collaborative operation of TPD effectively alleviates issues of excessive smoothing and artifact introduction, ensuring the integrity of spatial information across various network scales. Consequently, the proposed MambaNet significantly improves segmentation performance on noisy MRI data, especially for images with complex structural details.

In this paper, our contributions are as follows:

- We introduced a Poisson denoising module and applied it to MRI image denoising. This module effectively removes Poisson noise in MRI images and eliminates mixed noise, such as Gaussian and impulse noise.
- We designed the MambaNet network with a Twin-Path Decoder (TPD) branch structure, which reduces the loss of details in MRI images after denoising and enhances the model's ability to capture multi-level information.
- The segmentation performance of PDMambaNet was thoroughly validated on two public brain MRI datasets, with its effectiveness further confirmed by visualizing cortical thickness measurements.

II. METHOD

A. Architecture Overview

The PDMambaNet consists of two main components: the Poisson denoising module and the segmentation network. The noisy MRI image is first denoised by the Poisson denoising module. The denoised image is then fed into the segmentation network for feature extraction, and finally, it is segmented into the corresponding categories.

B. Possion Denoise

Inspired by the Poisson denoising scheme proposed in [5], we optimized the denoising process for noisy brain MRI images, enabling the Poisson denoising scheme to be smoothly transferred and applied to brain MRI image processing tasks. We made adaptive adjustments to the original scheme to suit the specific characteristics of brain MRI images. Specifically: Optimization of Denoising Algorithm Parameters: We optimized the parameters of the denoising algorithm to better adapt to low signal-to-noise ratios (SNR) and the complex anatomical structures of the brain. Incorporation of Medical Domain Prior Knowledge: We introduced constraints suitable for MRI images, ensuring that key brain details were preserved during the denoising process. These improvements enable the model to effectively reduce noise while preserving essential structural information, thereby significantly enhancing the segmentation accuracy of the model. The detailed derivation of the formulas can be found in the supplementary materials. Algorithm 1 illustrates the entire Poisson denoising process.

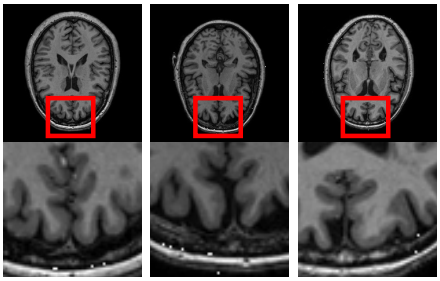
Algorithm 1: Poisson Denoise

Input: Noisy image X_0, number of iterations T
Output: Denoised image X^*
Initialize: Sparse representation $A^{(0)} = Encoder(X_0)$, iteration $t = 0$
while $t < T$ **do**
  Step 1: Update Image ;
  $X^{(t+1)} = Decoder(A^{(t)})$;
  Step 2: Compute Residual ;
  $R = X_0 - X^{(t+1)}$;
  Step 3: Update Sparse Representation ;
  $A^{(t+1)} = S(A^{(t)} + Encoder(R))$;
  Increment iteration: $t \leftarrow t + 1$;

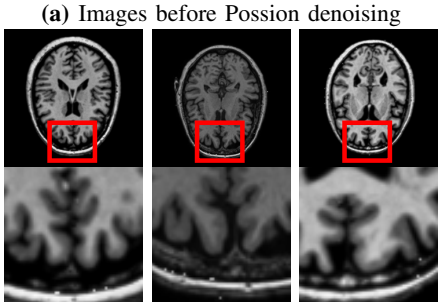
Fig. 2 compares results before and after Poisson denoising on the MRBrainS13 dataset. As shown in the zoomed-in view of the original image in Fig. 2 (a), the original image contains noise and has low contrast, which can lead to segmentation network errors. The zoomed-in view of the denoised result in Fig. 2 (b) demonstrates that Poisson denoising effectively removes noise. However, it also results in the loss of fine details in the MRI image, and the denoising process excessively smooths the image, causing edge blurring. This makes it challenging to distinguish crucial structural information, hindering the subsequent segmentation network's ability to utilize the multi-level information in the MRI image fully.

C. MambaNet

Fig. 3 shows the segmentation architecture of the proposed MambaNet, which is based on Mamba-UNet [28]. First, the input 2D grayscale image of size $H \times W \times 1$ is divided into patches and then into a one-dimensional sequence. This one-dimensional sequence is then adjusted to any dimension C through a linear embedding layer and processed through VSS blocks and block fusion layers. Each encoder layer in MambaNet extracts features using two VSS blocks, and the feature map sizes at each stage are $H/4 \times W/4$, $H/8 \times W/8$, $H/16 \times W/16$, and $H/32 \times W/32$, respectively.



(a) Images before Possion denoising



(b) Images after Poisson denoising

Fig. 2: Comparison of results before and after Poisson denoising on the MRBrainS13 dataset.

In MambaNet, we designed the TPD structure to better extract multi-level information from images processed by Poisson denoising. First, the multi-layer encoder extracts semantic information from the image while preserving fine features. One decoder in the TPD is used to recover the expression of fine features and restore the global structural information extracted by the multi-layer encoder. The final output is obtained by summing the outputs of the coarse decoder and the fine decoder. The feature maps of each stage of this decoder are of sizes $H/32 \times W/32$, $H/16 \times W/16$, $H/8 \times W/8$, and $H/4 \times W/4$, with the final segmentation map restored to $H \times W \times Class$ via linear projection.

The first two layers of the encoder, while not capturing deep semantic information and delicate features, contain low-level features such as edges and textures. The feature maps for these stages are $H/4 \times W/4$ and $H/8 \times W/8$. The second decoder in the TPD focuses on restoring features extracted by the first two layers of the encoder, effectively preserving details such as textures and edges. The feature maps for this decoder are $H/8 \times W/8$ and $H/4 \times W/4$, and the final coarse segmentation map, rich in low-level information, is restored to $H \times W \times Class$ via linear projection. This complementary relationship between low-level and high-level information enables the model to reduce the loss of details and enhance its capability to capture multi-level information. The collaborative operation of the TPD effectively mitigates the issues of over-smoothing and artifact introduction caused by Poisson denoising, ensuring the integrity of spatial information across various network scales. The bottleneck layer of MambaNet consists of two VSS blocks. However, in the coarse decoder, we did not use the bottleneck layer; instead, the downsampled features were directly passed to the decoder to prevent over-decoding.

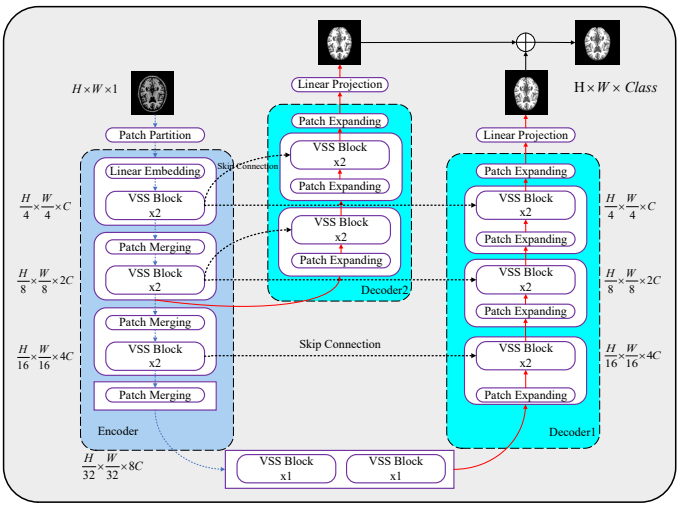


Fig. 3: Architecture diagram of PDMambaNet

III. Experiments and Results

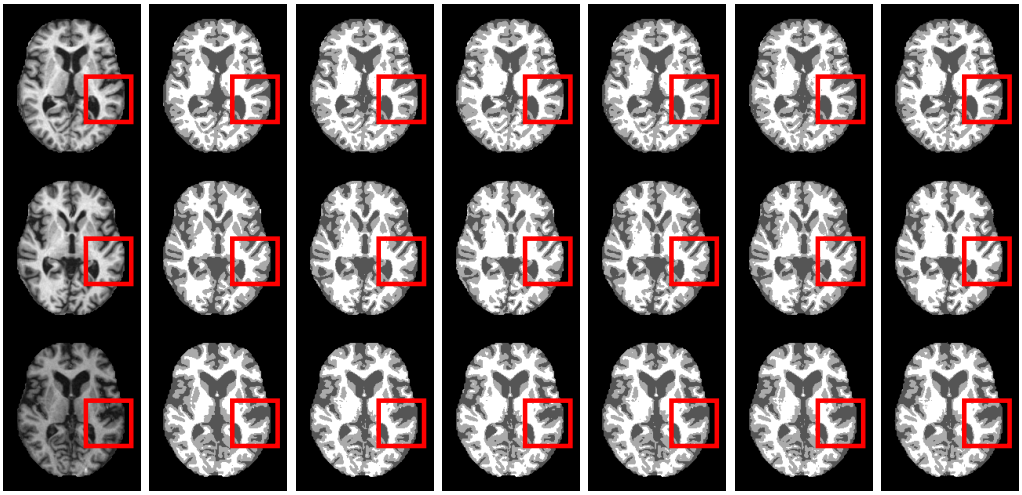
A. Datasets

1) OASIS1: The OASIS-1 dataset [33] used in this experiment is from the Open Access Series of Imaging Studies (OASIS). It comprises 421 subjects aged between 18 and 96 years. Each subject has a T1-weighted magnetic resonance imaging (MRI) scan. The dataset labels classify brain tissue into the cerebrospinal fluid (CSF), grey matter (GM), and white matter (WM).

2) MRBrainS13: The MRBrainS13 challenge dataset consists of 20 subjects acquired using a 3.0T Philips Achieva MR scanner at the University Medical Center Utrecht, Netherlands [34]. The dataset includes multi-sequence MRI brain scans, such as T1, T1-IR and T2-FLAIR used for the challenge. The dataset labels classify brain tissue into the cerebrospinal fluid (CSF), grey matter (GM), and white matter (WM).

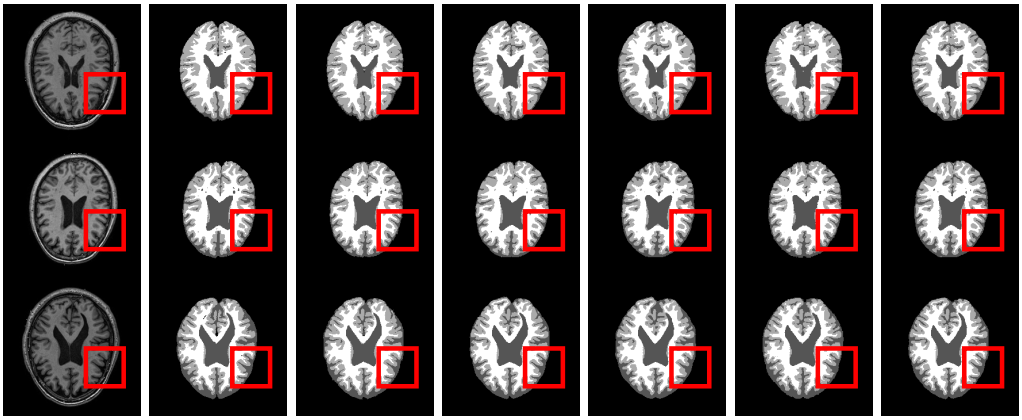
B. Implementation Details

The experiments were conducted on an Ubuntu 22.04 system equipped with an Nvidia A40 GPU with 48GB of memory, using Python 3.8.19, PyTorch 2.2.0, and CUDA 11.8. The model was used for 2D medical image segmentation. We randomly divided the two datasets into training, test, and validation sets in a ratio of 8:1:1. All images were normalized and resized to 224×224. Data augmentation techniques, including vertical flipping, horizontal flipping, and random rotation, were applied. PDMambaNet was trained for 40,000 iterations on the OASIS1 dataset and 20,000 iterations on the MRBrainS13 dataset, with a batch size of 12. The stochastic gradient descent (SGD) optimizer [35] was used with a learning rate of 0.01, a momentum of 0.9, and a weight decay of 0.0001. The network performance was evaluated on the validation set every 200 iterations, and the model weights were saved only when the validation set achieved new best performance.



(a) (b) (c) (d) (e) (f) (g)

Fig. 4: Segmentation results comparison among different models on the OASIS1 dataset. (a) Image, (b) GT, (c) Mamba-UNet, (d) UNet, (e) PNet, (f) Swin-UNet, (g) PDMambaNet.



(a) (b) (c) (d) (e) (f) (g)

Fig. 5: Segmentation results comparison among different models on the MRBrainS13 dataset: (a) Image, (b) GT, (c) Mamba-UNet, (d) UNet, (e) PNet, (f) Swin-UNet, (g) PDMambaNet.

C. Qualitative Results

As shown in Fig.4 and Fig.5, PDMambaNet can more accurately identify and segment each category compared to other algorithms. Additionally, the zoomed-in views in Fig.4 demonstrate that PDMambaNet excels in recognizing fine structural details when segmenting complex structures, such as edges and textures between different tissues. Similarly, the zoomed-in views in Fig.5 illustrate that PDMambaNet effectively mitigates the edge over-smoothing issue caused by Poisson denoising while still accurately identifying the edge structures of the image.

D. Quantitative Results

To ensure a fair comparison, Mamba-UNet [28], UNet [10], Swin-UNet [14], and PNet [36] were all trained under the same hyperparameter configurations and directly compared with PDMambaNet. Seven objective evaluation metrics were used to quantitatively compare the proposed method: 1) Similarity measures, including Dice, Accuracy (Acc), Precision (Pre), Sensitivity (Sen), and Specificity (Spe), indicated by an upward arrow (↑), where values closer to 1 indicate better performance. 2) Discrepancy measures, including the 95% Hausdorff Distance (HD95) and Average Surface Distance (ASD), indicated by a downward arrow (↓), where lower values are better, representing a higher similarity between the predicted segmentation and the ground truth segmentation. Table I and Table II present the average evaluation metrics, including similarity and dissimilarity measures, for PDMambaNet and other segmentation networks on the OASIS1 and MRBrainS13 datasets. The best results are highlighted in bold. The quantitative analysis shows that PDMambaNet consistently achieves high segmentation accuracy on large-scale datasets and precisely predicts segmentation masks on small-scale datasets. The high Acc and Sen values indicate the model's accuracy, suggesting that the introduction of Poisson denoising effectively removes noise from MRI images. Moreover, the high Dice and HD95 scores

TABLE I: Average segmentation metrics for different segmentation networks on the OASIS1 dataset.

Model	Dice(↑)	Acc(↑)	Pre(↑)	Sen(↑)	Spe(↑)	HD95(↓)	ASD(↓)
Mamba-UNet [28]	0.9097	0.9739	0.9104	0.9191	0.9783	2.1684	0.6380
UNet [10]	0.8743	0.9678	0.8907	0.8807	0.9745	2.4922	0.7795
PNet [36]	0.9262	0.9757	0.9261	0.9227	0.9798	1.6571	**0.3753**
Swin-UNet [14]	0.9248	0.9728	**0.9338**	0.9242	**0.9810**	**1.4741**	0.3918
PDMambaNet	**0.9266**	**0.9773**	0.9196	**0.9300**	**0.9810**	1.6410	0.4700

TABLE II: Average segmentation metrics for different segmentation networks on the MRBrainS13 dataset.

Model	Dice(↑)	Acc(↑)	Pre(↑)	Sen(↑)	Spe(↑)	HD95(↓)	ASD(↓)
Mamba-UNet [28]	0.6908	0.8194	0.6940	0.7098	0.8264	3.3607	0.9981
UNet [10]	0.7048	0.8214	0.7093	0.7212	0.8281	3.2435	0.9775
PNet [36]	0.7068	0.8232	0.7052	0.7236	0.8251	2.9803	0.8765
Swin-UNet [14]	0.7102	0.8312	**0.7152**	0.7211	0.8385	**2.2916**	**0.7697**
PDMambaNet	**0.7158**	**0.8319**	0.7045	**0.7352**	**0.8412**	3.0223	0.7706

demonstrate that PDMambaNet efficiently leverages multi-level information from MRI images, and the collaborative optimization of this information enhances the model's feature representation capabilities. The consistent performance across datasets of varying scales further highlights the robustness and generalization ability of PDMambaNet.

E. Ablation Studies and Analysis

The PDMambaNet involves two key components: 1) Poisson denoising and 2) Twin-Path Decoder (TPD). We conducted ablation studies to compare the proposed components of our algorithm. We performed identical ablation experiments on both datasets to validate the proposed model's effectiveness and its improvements. The results on the OASIS1 dataset are shown in Table III, and the results on the MRBrainS13 dataset are shown in Table IV. TPD represents the Twin-Path Decoder branch structure in these tables, and PD represents the Poisson denoising component. A ✓ indicates that the component was used, and an ✗ indicates that the component was not used.

TABLE III: Ablation results on the OASIS1 dataset.

PD	TPD	Dice(↑)	Acc(↑)	Pre(↑)	Sen(↑)	Spe(↑)	HD95(↓)	ASD(↓)
✗	✗	0.9097	0.9739	0.9104	0.9191	0.9783	2.1684	0.6380
✓	✗	0.9064	0.9740	0.8983	0.9160	0.9790	1.9387	0.5929
✗	✓	0.9170	0.9757	**0.9232**	0.9206	0.9799	1.6775	0.4760
✓	✓	**0.9266**	**0.9773**	0.9196	**0.9300**	**0.9810**	**1.6410**	**0.4700**

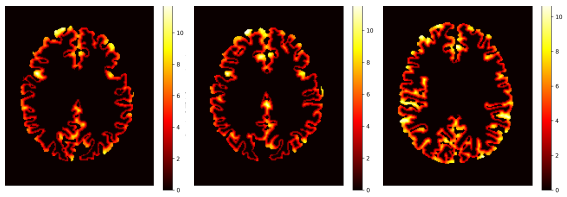
TABLE IV: Ablation results on the MRBrainS13 dataset.

PD	TPD	Dice(↑)	Acc(↑)	Pre(↑)	Sen(↑)	Spe(↑)	HD95(↓	ASD(↓)
✗	✗	0.6908	0.8194	0.6940	0.7098	0.8264	3.3607	0.9981
✓	✗	0.6957	0.8235	0.7015	0.7095	0.8308	3.4576	1.1198
✗	✓	0.7068	0.8175	0.7014	0.7239	0.8241	**2.5121**	**0.7429**
✓	✓	**0.7158**	**0.8319**	**0.7045**	**0.7352**	**0.8412**	3.0223	0.7706

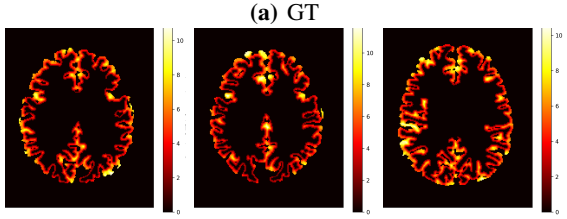
We conducted an objective evaluation of the proposed components. The results in Table III and Table IV indicate that while Poisson denoising effectively removes noise and improves segmentation accuracy, the HD95 and ASD metrics perform poorly. This suggests that Poisson denoising leads to the loss of fine details and edge blurring in MRI images. To further validate the role of TPD, the results of using TPD alone in this section demonstrate that TPD reduces detail loss and enhances the model's ability to express multi-level information. Additionally, the collaborative work between TPDs effectively mitigates the issues of excessive image smoothing and artefact introduction, ensuring the maintenance of spatial information across various network scales. PDMambaNet combines the advantages of both components, effectively removing noise from MRI images while maintaining high segmentation accuracy.

F. Cortical Thickness Measurement Results

As shown in the cortical thickness measurement results in Fig. 6, PDMambaNet's segmentation results are highly consistent with the ground truth (GT). This consistency validates the effectiveness of PDMambaNet in segmentation tasks and enhances the accuracy of subsequent cortical thickness measurements. By accurately delineating the boundaries between gray matter and other tissue types, PDMambaNet reliably reveals subtle variations in cortical thickness across different brain regions, aiding in the study of cortical complexity. This capability is particularly valuable for investigating morphological changes associated with neurological diseases and identifying minor cortical thickness variations in clinical settings.



(a) GT



(b) PDMambaNet

Fig. 6: Cortical Thickness Measurement Map(mm) of Segmentation Results from the OASIS1 Dataset.

IV. Conclusion

This paper proposes PDMambaNet, which addresses noise removal in MRI images through a Poisson denoising module. However, Poisson denoising often leads to the loss of important brain tissue details and blurred edges, causing false positives and false negatives in downstream segmentation tasks. To overcome this, we introduce a segmentation network with a Two-Path Decoder (TPD) architecture, which reduces the loss of details and edge blurring caused by over-smoothing, while effectively minimizing artifact introduction. This ensures high segmentation accuracy.

REFERENCES

[1] Zhaolin Chen, Kamlesh Pawar, Mevan Ekanayake, Cameron Pain, Shenjun Zhong, and Gary F Egan, "Deep learning for image enhancement and correction in magnetic resonance imaging—state-of-the-art and challenges," *Journal of Digital Imaging*, vol. 36, no. 1, pp. 204–230, 2023.

[2] Liangyu Chen, Xin Lu, Jie Zhang, Xiaojie Chu, and Chengpeng Chen, "Hinet: Half instance normalization network for image restoration," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 182–192.

[3] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao, "Multi-stage progressive image restoration," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 14821–14831.

[4] Samuel W Hasinoff, "Photon, poisson noise," in *Computer vision: a reference guide*, pp. 980–982. Springer, 2021.

[5] Calvin-Khang Ta, Abhishek Aich, Akash Gupta, and Amit K Roy-Chowdhury, "Poisson2sparse: Self-supervised poisson denoising from a single image," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2022, pp. 557–567.

[6] James Pawley, "Fundamental limits in confocal microscopy," in *Handbook of biological confocal microscopy*, pp. 15–26. Springer, 2006.

[7] Lucas Sjulson and Gero Miesenbock, "Optical recording of action potentials and other discrete physiological events: a perspective from signal detection theory," *Physiology*, vol. 22, no. 1, pp. 47–55, 2007.

[8] Quande Liu, Cheng Chen, Jing Qin, Qi Dou, and Pheng-Ann Heng, "Feddg: Federated domain generalization on medical image segmentation via episodic learning in continuous frequency space," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 1013–1023.

[9] Feifei Zhang, Fei Shi, Dayong Ren, and Yue Li, "A fuzzy c-means clustering algorithm for real medical image segmentation," in *ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2025, pp. 1–5.

[10] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*. Springer, 2015, pp. 234–241.

[11] Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang, "Unet++: A nested u-net architecture for medical image segmentation," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4*. Springer, 2018, pp. 3–11.

[12] Huimin Huang, Lanfen Lin, Ruofeng Tong, Hongjie Hu, Qiaowei Zhang, Yutaro Iwamoto, Xianhua Han, Yen-Wei Chen, and Jian Wu, "Unet 3+: A full-scale connected unet for medical image segmentation," in *ICASSP 2020-2020 IEEE international conference on acoustics, speech and signal processing (ICASSP)*. IEEE, 2020, pp. 1055–1059.

[13] Jieneng Chen, Yongyi Lu, Qihang Yu, Xiangde Luo, Ehsan Adeli, Yan Wang, Le Lu, Alan L Yuille, and Yuyin Zhou, "Transunet: Transformers make strong encoders for medical image segmentation," *arXiv preprint arXiv:2102.04306*, 2021.

[14] Hu Cao, Yueyue Wang, Joy Chen, Dongsheng Jiang, Xiaopeng Zhang, Qi Tian, and Manning Wang, "Swin-unet: Unet-like pure transformer for medical image segmentation," in *European conference on computer vision*. Springer, 2022, pp. 205–218.

[15] Ailiang Lin, Bingzhi Chen, Jiayu Xu, Zheng Zhang, Guangming Lu, and David Zhang, "Ds-transunet: Dual swin transformer u-net for medical image segmentation," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–15, 2022.

[16] Yundong Zhang, Huiye Liu, and Qiang Hu, "Transfuse: Fusing transformers and cnns for medical image segmentation," in *Medical image computing and computer assisted intervention–MICCAI 2021: 24th international conference, Strasbourg, France, September 27–October 1, 2021, proceedings, Part I 24*. Springer, 2021, pp. 14–24.

[17] Jonathan Long, Evan Shelhamer, and Trevor Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.

[18] Dayong Ren, Zhengyi Wu, Jiawei Li, Piaopiao Yu, Jie Guo, Mingqiang Wei, and Yanwen Guo, "Point attention network for point cloud semantic segmentation," *Science China Information Sciences*, vol. 65, no. 9, pp. 192104, 2022.

[19] Dayong Ren, Zhe Ma, Yuanpei Chen, Weihang Peng, Xiaode Liu, Yuhan Zhang, and Yufei Guo, "Spiking pointnet: Spiking neural networks for point clouds," *Advances in Neural Information Processing Systems*, vol. 36, 2024.

[20] Dayong Ren, Jiawei Li, Zhengyi Wu, Jie Guo, Mingqiang Wei, and Yanwen Guo, "Mffnet: multimodal feature fusion network for point cloud semantic segmentation," *The Visual Computer*, vol. 40, no. 8, pp. 5155–5167, 2024.

[21] Yanwen Guo, Yuanqi Li, Dayong Ren, Xiaohong Zhang, Jiawei Li, Liang Pu, Changfeng Ma, Xiaoyu Zhan, Jie Guo, Mingqiang Wei, et al., "Lidar-net: A real-scanned 3d point cloud dataset for indoor scenes," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 21989–21999.

[22] Chen Chen, Yisen Wang, Honghua Chen, Xuefeng Yan, Dayong Ren, Yanwen Guo, Haoran Xie, Fu Lee Wang, and Mingqiang Wei, "Geosegnet: point cloud semantic segmentation via geometric encoder–decoder modeling," *The Visual Computer*, vol. 40, no. 8, pp. 5107–5121, 2024.

[23] Dayong Ren, Shuangyu Yang, WEN JIE LI, Jie Guo, and Yanwen Guo, "Sae: Estimation for transition matrix in annotation algorithms," .

[24] Harsh Mehta, Ankit Gupta, Ashok Cutkosky, and Behnam Neyshabur, "Long range language modeling via gated state spaces," *arXiv preprint arXiv:2206.13947*, 2022.

[25] Albert Gu and Tri Dao, "Mamba: Linear-time sequence modeling with selective state spaces," *arXiv preprint arXiv:2312.00752*, 2023.

[26] Jun Ma, Feifei Li, and Bo Wang, "U-mamba: Enhancing long-range dependency for biomedical image segmentation," *arXiv preprint arXiv:2401.04722*, 2024.

[27] Lianghui Zhu, Bencheng Liao, Qian Zhang, Xinlong Wang, Wenyu Liu, and Xinggang Wang, "Vision mamba: Efficient visual representation learning with bidirectional state space model," *arXiv preprint arXiv:2401.09417*, 2024.

[28] Ziyang Wang, Jian-Qing Zheng, Yichi Zhang, Ge Cui, and Lei Li, "Mamba-unet: Unet-like pure visual mamba for medical image segmentation," *arXiv preprint arXiv:2402.05079*, 2024.

[29] Weibin Liao, Yinghao Zhu, Xinyuan Wang, Cehngwei Pan, Yasha Wang, and Liantao Ma, "Lightm-unet: Mamba assists in lightweight unet for medical image segmentation," *arXiv preprint arXiv:2403.05246*, 2024.

[30] Renkai Wu, Yinghao Liu, Pengchen Liang, and Qing Chang, "Ultralight vm-unet: Parallel vision mamba significantly reduces parameters for skin lesion segmentation," *arXiv preprint arXiv:2403.20035*, 2024.

[31] Feifei Zhang, Fei Shi, Dayong Ren, Zhenhong Jia, and Jianyi Wang, "Dual-mambanet: A lightweight dual-branch brain image segmentation network based on local attention and mamba," in *International Conference on Pattern Recognition*. Springer, 2024, pp. 92–107.

[32] Md Maklachur Rahman, Abdullah Aman Tutul, Ankur Nath, Lamyanba Laishram, Soon Ki Jung, and Tracy Hammond, "Mamba in vision: A comprehensive survey of techniques and applications," *arXiv preprint arXiv:2410.03105*, 2024.

[33] Daniel S Marcus, Tracy H Wang, Jamie Parker, John G Csernansky, John C Morris, and Randy L Buckner, "Open access series of imaging studies (oasis): cross-sectional mri data in young, middle aged, nondemented, and demented older adults," *Journal of cognitive neuroscience*, vol. 19, no. 9, pp. 1498–1507, 2007.

[34] Adriënne M Mendrik, Koen L Vincken, Hugo J Kuijf, Marcel Breeuwer, Willem H Bouvy, Jeroen De Bresser, Amir Alansary, Marleen De Bruijne, Aaron Carass, Ayman El-Baz, et al., "Mrbrains challenge: online evaluation framework for brain image segmentation in 3t mri scans," *Computational intelligence and neuroscience*, vol. 2015, no. 1, pp. 813696, 2015.

[35] Léon Bottou et al., "Stochastic gradient learning in neural networks," *Proceedings of Neuro-Nımes*, vol. 91, no. 8, pp. 12, 1991.

[36] Guotai Wang, Maria A Zuluaga, Wenqi Li, Rosalind Pratt, Premal A Patel, Michael Aertsen, Tom Doel, Anna L David, Jan Deprest, Sébastien Ourselin, et al., "Deepigeos: a deep interactive geodesic framework for medical image segmentation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 41, no. 7, pp. 1559–1572, 2018.