ResMAP: Restoring MRIs of Mixed Artifacts by Prompt Cascading Retrieval

Yuxian Tang[1†], Feng Li[1†], Feng Shi[2], and Qian Wang[1,3(✉)]

[1] School of Biomedical Engineering & State Key Laboratory of Advanced Medical Materials and Devices, ShanghaiTech University, Shanghai, China
qianwang@shanghaitech.edu.cn
[2] Department of Research and Development, United Imaging Intelligence, Shanghai, China
[3] Shanghai Clinical Research and Trial Center, Shanghai, China

Abstract. Image Restoration (IR) aims to enhance degraded images to provide high-quality diagnostic references in Magnetic Resonance Imaging (MRI). Although recent All-in-One IR (AiOIR) methods seek to handle multiple artifacts within a unified network, they still struggle with mixed artifacts, where multiple unknown artifacts occur simultaneously in a single MRI scan. To tackle this challenge, we propose ResMAP, a cascading framework for **Res**toring MRIs of **M**ixed **A**rtifacts by **P**rompt Retrieval. It is trained exclusively on individual artifact types but can effectively handle all their mixed forms in inference, offering a feasible solution instead of requiring exhaustive training on mixed artifacts. Specifically, our ResMAP utilizes a coarse-to-fine correction process for mixed artifacts by cascading retrieval of prompts based on the artifact types. In this process, the retrieval guidance is provided through the perception and classification of fine-grained image features, while the prompts are prepared via LLM-based generation and fine-tuning. Validations on three types of artifacts and their mixed forms demonstrate the superiority of ResMAP over current IR methods. Besides, zero-shot experiments on MRIs from multiple field strengths further confirm the promising generalizability of the proposed framework. Our code is available at https://github.com/Tanishabc/ResMAP.

Keywords: Brain MRI · Mixed Artifacts · All-in-one Restoration · Cross-modal Prompt.

1 Introduction

Magnetic Resonance Imaging (MRI) is essential in medical diagnosis due to its non-invasive nature and ability to provide detailed anatomical insights [1]. However, the prolonged imaging process often leads to inevitable subject movements, while unexpected hardware imperfections in the scanner can also compromise image quality [19]. In this context, Image Restoration (IR) [2] aims

† These authors contributed equally to this work.

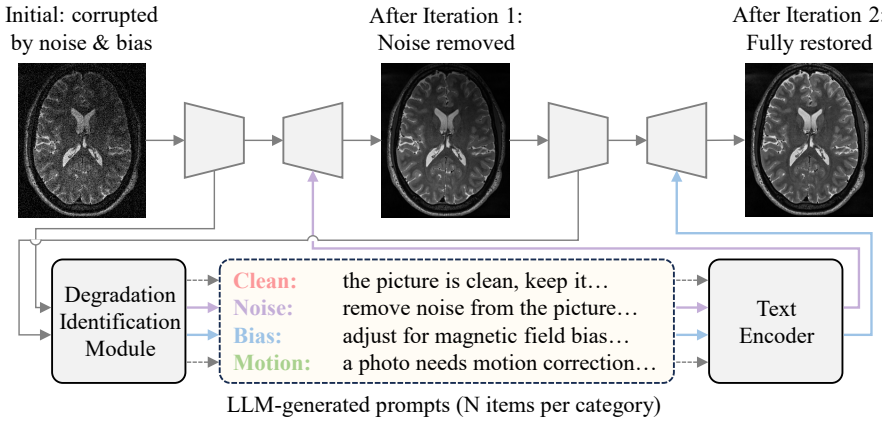


Fig. 1. An example of brain MRI with mixed degradation (e.g., noise and bias) is restored iteratively based on degradation identification. The LLM-generated prompt is picked and fused with the original image feature for degradation-aware restoration in individual iteration.

to enhance low-quality images retrospectively, mitigating the adverse impact of quality degradation. Within the scope of single-task, IR has achieved remarkable success such as motion correction [6], noise removal [12], and bias field correction [9]. However, these methods lack generalization beyond the specific task, limiting their ability in clinical scenarios where multiple artifacts often occur simultaneously [11].

With advances in multi-task deep learning, All-in-One Image Restoration (AiOIR) [10] endeavors to handle multiple image enhancement tasks within a single model, including those involving multiple artifacts and multiple modalities. It employs various techniques to adapt the shared network to different tasks, including multimodal approaches [4], prompt learning [13], and Mixture of Experts (MoE) [16]. For instance, Airnet [14] deploys an additional encoder built upon the backbone restoration network to simultaneously process several types of image corruption: denoising, deraining, and dehazing, while PromptIR [18] introduces a plug-in module that can be integrated into existing restoration networks for dynamic domain adaptation to the different degradation patterns. In the field of medical imaging, a pioneering work of AMIR [23] adopts a task-adaptive routing strategy based on MoE to restore major modalities (MRI, CT, and PET). Despite upgrading the models' utility from single to multiple tasks, they are still limited to correcting the degradations covered during training rather than addressing complex unseen mixed degradations. Considering the common yet unresolved problem of mixed artifact correction appeared in a single MR scan, Jiang et al. [11] took two types of mixed-artifact MR images (motion mixed with low resolution and motion mixed with noise) for model development. However, such an intuitive training strategy is inefficient for covering all artifact combination possibilities, especially as the number of artifact types increases, the difficulty

of training increases exponentially. Given the numerous types of degradations in MRI [5], training the model on all possible combinations is impractical.

To address these limitations, we introduce a cascading restoration framework ResMAP to adaptively remove mixed artifacts present in an MRI image, where a degradation-aware prompt guides the sequence of each artifact removal. Notably, ResMAP is only trained on several types of individual artifacts but can be used directly for their mixtures during inference, which makes the effective removal of various mixed artifacts possible. For example, a brain MRI image corrupted by noise and bias is restored through two cascaded restoration networks as shown in Fig. 1. To enable the same restoration network to dynamically adapt to varying artifact inputs, artifact identification is first performed based on the encoded image features. Subsequently, LLM-generated prompts corresponding to the identified results are sampled and encoded, providing rich semantic guidance for artifact-specific correction. We conducted experiments to validate the proposed method on three types of artifacts and their mixtures. The results demonstrate that ResMAP outperforms existing approaches. Zero-shot validation on MRI data from different field strengths further shows its promising applicability. Our main contributions can be summarized as follows:

1. We address MRI mixed-artifact removal through a cascading architecture guided by prompts, extending existing AiOIR methods beyond their limitation of handling only predefined artifact types.

2. We propose a text prompt retrieval strategy based on image artifact identification, enabling degradation-aware guidance for dynamic adaptive correction in AiOIR and mixed artifact scenarios.

2 Method

Fig. 2 depicts an overview of the proposed framework for image restoration based on artifact-specific prompt retrieval, structured into three key modules. Module (a) performs degradation classification and subsequently utilizes the predicted result for prompt retrieval. In Module (b), separate training is conducted to prepare the prompt embedding pool. Module (c) illustrates the fusion process of the selected prompt with the image feature.

2.1 Network Architecture and Training Scheme

Network Architecture We construct our ResMAP using Restormer [24] as the backbone, which is an efficient Unet-style image restoration model. The core innovation of ResMAP lies in the integration of artifact-specific prompts to facilitate multi-task learning and enable effective mixed artifact removal. Specifically, for a given degraded input image I_{LQ}, the encoder $\mathcal{E}_I$ of the restoration network extracts the image feature $\mathcal{F}_i$, which is fed into a classifier $\mathcal{C}_I$ to predict the degradation category $\hat{y}_i = \mathcal{C}_I(\mathcal{E}_I(I_{LQ}))$. Based on this prediction, a text prompt feature $\mathcal{F}_t$ is sampled from $\mathcal{F}_{pi}$, a matching prompt group with the same class label $\hat{y}_i$. Then, the selected $\mathcal{F}_t$ is integrated with the image feature $\mathcal{F}_i$ through a

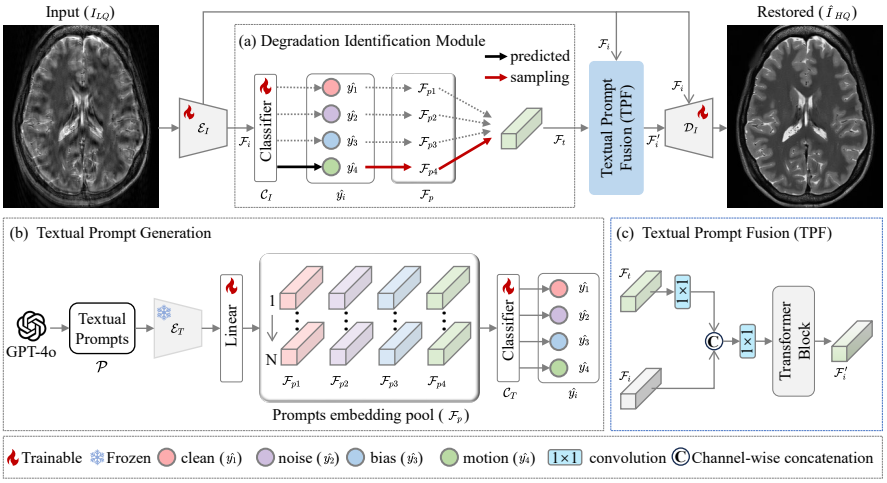


Fig. 2. The overview of the proposed ResMAP. (a) Degradation Identification Module for artifact removal prompt retrieval. (b) LLM-based prompts generation and encoding. (c) Image-Text feature Fusion based on Transformer Block.

Textual Prompt Fusion (TPF) module. Finally, the decoder progressively reconstructs a high-resolution output $\hat{I}_{HQ}$ by processing the original image feature $\mathcal{F}_i$ with the guidance of fused feature $\mathcal{F}_i'$. The preparation of the prompt embedding pool $\mathcal{F}_p$ and the TPF module will be detailed in Section 2.2.

Training Scheme In our framework, artifact detection and image correction are jointly trained. Specifically, we develop a two-layer MLP classifier $\mathcal{C}_I$ to predict artifact types: $\hat{y}_i = \mathcal{C}_I(\mathcal{F}_i)$, where $\hat{y}_i \in \mathbb{R}^D$, with D representing degradation categories plus one for clean images. We employ cross-entropy loss $\mathcal{L}_{class}$ for multi-artifact classification. For image restoration, ℓ_1 loss is used to minimize the pixel-wise difference between the corrected image $\hat{I}_{HQ}$ and the ground truth I_{HQ}. The overall training loss is formulated as:

$$\mathcal{L}_{total} = ||I_{HQ} - \hat{I}_{HQ}||_{\ell_1} + \lambda \mathcal{L}_{class} \tag{1}$$

the weighting coefficient λ balancing the two loss terms is set to 0.01.

2.2 Prompt Generation and Fusion

Prompt Generation Inspired by InstructIR [3], we leverage LLM-generated prompts for multi-task image restoration. Specifically, GPT-4o generates N prompts per class (motion, noise, bias, clean), forming a pool of 4N (N = 1000). These prompts, denoted as $\mathcal{P}$, are processed by a frozen text encoder $\mathcal{E}_T$ to obtain 384-dimensional feature vectors. We adopt bge-micro-v2, a distilled BGE-SMALL-EN [22] text encoder, as in InstructIR. To enhance intra-category compactness,

features are fine-tuned via a linear layer: $\mathcal{F}_p = \mathbb{L}(\mathcal{E}_T(\mathcal{P}))$, then classified by a lightweight $\mathcal{C}_T$ (same as $\mathcal{C}_I$). Fine-tuning uses cross-entropy loss, and the refined prompt features are readily retrievable for image correction.

Textual Prompt Fusion (TPF) We perform image-text feature fusion as:

$$\mathcal{F}_i' = Block(Conv[\mathcal{F}_i; Conv(\mathcal{F}_t)] \tag{2}$$

where $\mathcal{F}_i$ is the input image feature and $\mathcal{F}_i'$ is the fused version, with [;] denoting concatenation and $Conv$ as a 1×1 convolution for feature alignment in channel dimension. The Block is a standard Transformer Block from Restormer [24]. The TPF is incorporated between decoder layers to exploit cross-model information.

2.3 Inference: Automatic Cascading Prompts for Mixed Artifacts

As training on mixed artifacts is impractical to scale, ResMAP is trained on individual artifacts and applied to their mixtures in inference. As shown in Fig. 1, an image corrupted by noise and bias is first identified as noise, and a prompt from the same group is picked and fused into the decoder of the restoration network, enable it perform noise removal. Subsequently, the denoised image is identified as bias and undergoes corresponding correction. We will demonstrate the correction order decision and its superiority in subsequent experiments. The overall process terminates when the degradation module outputs clean, with a maximum of three iterations to prevent over-correction and to information loss.

3 Experiments and Results

3.1 Dataset and Experimental Setup

Dataset. We synthesized motion, noise, and bias field artifacts from T2-weighted MRI scans of 120 HCP subjects, resulting in four image types, including clean images. For each type, 50 axial slices per subject yielded 20,000 training slices (100 subjects) and 4,000 validation slices (20 subjects). Motion artifacts were generated as in [7], simulating 3D k-space translations and rotations with parameters from $\mathcal{N}(0, 10)$. Noise artifacts were simulated using TorchIO [17] as

Table 1. Comparison of ResMAP with other AiOIR methods on separate artifacts.

Method	Noise		Bias		Motion		Average	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Airnet [14]	38.23	0.971	36.89	0.979	32.16	0.894	35.76	0.948
PromptIR [18]	38.69	0.973	37.35	0.982	32.35	0.919	36.13	0.958
AMIR [23]	38.72	0.974	37.52	0.984	32.24	0.920	36.16	0.959
ResMAP (Ours)	**38.84**	**0.976**	**37.84**	**0.987**	**33.12**	**0.963**	**36.60**	**0.975**
ResMAP (ACC)	0.951		0.955		0.962		0.956	

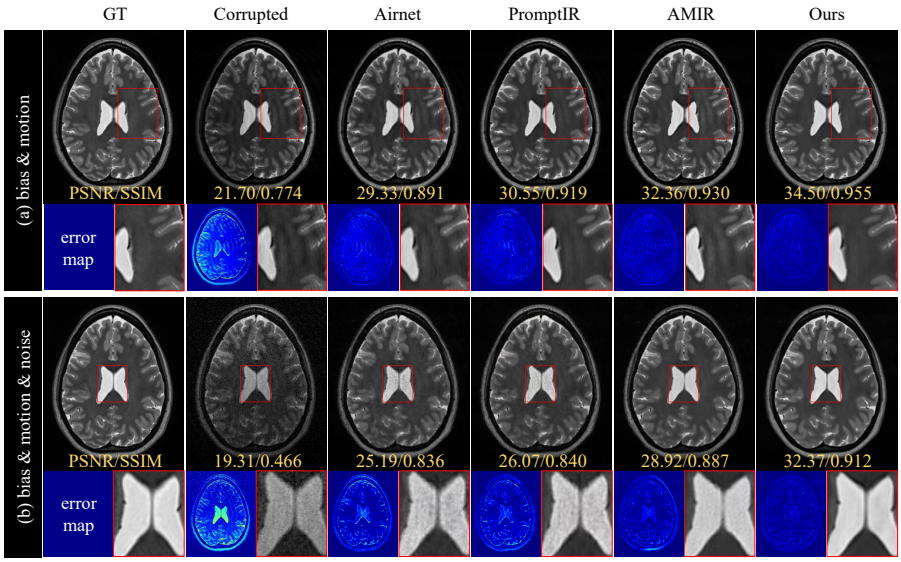


Fig. 3. Comparison of ResMAP with other methods on mixed artifacts. Bright and dark blue in error maps represent higher and lower correction deviation, respectively.

in [20], with $\sigma_{noise} = [0, 60]$ and $\sigma_{blur} = [0, 1]$. Bias fields were modeled via polynomial fitting as in [21] and implemented by NiftyNet [8], with order $\in$ [1,3] and coefficient $\in$ [0,1]. To evaluate mixed artifact removal, four artifact combinations were applied to each of 10 additional HCP subjects, independent of the previous 120. For zero-shot evaluation across field strengths, we used 10 subjects each from the 0.064T LISA challenge 2024 and 0.3T M4RAW [15], as well as 3 healthy subjects from our in-house 5T MRI dataset (TE = 271.46 ms, TR = 2 s, slice thickness = 0.6 mm, and FOV = 0.59 $\times$ 0.59 mm^2).

Implementation Details. The model is trained on 64 $\times$ 64 image patches using the Adam optimizer with a learning rate of 2×10^4 , a batch size of 32 over 25k iterations. All experiments are implemented in PyTorch and conducted on an NVIDIA A100 GPU with 40GB memory. Evaluation metrics include PSNR, SSIM, absolute error maps, and qualitative comparisons of corrected results.

3.2 Comparison with State-of-the-Art Methods

All-in-One We first compare ResMAP with other AiOIR methods for removing three independent types of artifacts using a unified network. Table 1 summarizes the restoration results and provides the classification accuracy (ACC) of ResMAP for different artifacts. Our method consistently outperforms all baselines, demonstrating the effectiveness of feature embeddings extracted from artifact-specific textual descriptions in improving image restoration performance.

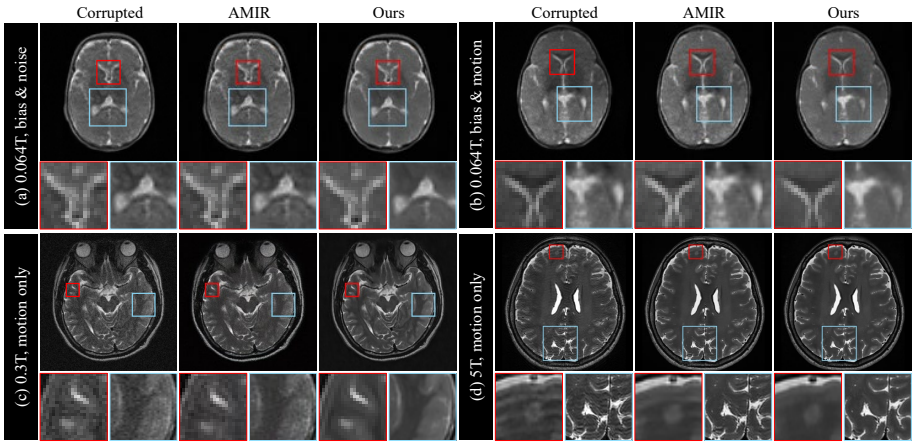


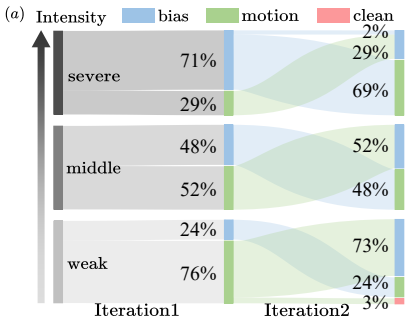
Fig. 4. Comparison of ResMAP with AMIR on brain MRI from multiple field strengths.

Mixed Artifacts Fig. 3 shows qualitative and quantitative comparisons of representative cases with mixed artifacts. To be fair, all models were iterated according to the number of artifacts, e.g., iterating twice for case (a) with bias and motion mixed. Despite under the same numbers of iterations, our method achieves the lowest correction error for mixtures of both two or three artifact categories. This confirms that the artifact-specific cross-model prompt retrieval outperforms other image-only based adaptive approaches.

Zero-shot We further compare ResMAP with AMIR, the second-best method, across three external datasets. The 0.064T LISA dataset, a very-low-field MRI set with real mixed artifacts, is visualized in Fig. 4 (case a and b). The 0.3T M4RAW and the 5T in-house dataset contain real and synthetic motion artifacts, respectively, as shown in cases (c) and (d). All results demonstrate ResMAP outperforms AMIR in both artifact removal and detail restoration.

Table 2. Ablation study of the key components on both single and mixed artifact.

Experimental Setting				Metric	
artifact	classification	Prompt	Cascading	PSNR	SSIM
Single	×	×	×	35.94	0.956
	✓	×	×	36.47	0.968
	✓	✓	×	36.60	0.975
Mixed	×	×	×	20.06	0.591
	×	×	✓	22.73	0.682
	✓	✓	✓	24.39	0.775



(b)

Intensity	Order	PSNR	SSIM
severe	Inference	23.08	0.743
	Reverse	20.39	0.611
middle	Inference	24.41	0.789
	Reverse	22.29	0.633
weak	Inference	25.68	0.793
	Reverse	21.63	0.625

Fig. 5. (a) Correction order of MRIs with mixed artifacts across three bias intensities. (b) Restoration comparison between inference and its reverse order.

3.3 Ablation Study

We conducted ablation studies on two settings: all-in-one and mixed artifacts. **In the all-in-one setting**, we assessed the impact of degradation classification and LLM-generated prompts, with metrics averaged over three artifact types. As shown in Table 2, the addition of the classification significantly improves baseline performance, while LLM-generated prompts further enhance the model. **In the mixed artifacts setting**, we demonstrate that, the incorporation of the cascaded architecture enables ResMAP to achieve significantly better performance than the baseline model.

3.4 Effectiveness in Degradation-Aware Module

To correct an MR scan with mixed artifacts, we design it as a cascading network where the forward path relies on appropriate prompt retrieval, with degradation awareness playing a key role. We demonstrate the processing mechanism for adaptive correction order of the cascaded architecture and prove its superiority through two experiments. First, we added fixed motion and varying bias field artifacts (weak, middle, and severe) to three image groups, then performed artifact detection and correction using two cascaded ResMAP. As shown in Fig. 5 A, the proportion of bias field artifacts detected in the first iteration gradually increases with the intensity of the bias field. In the second stage, the detection results are almost the opposite of the first stage (images initially detected as bias field are corrected and then detected as motion to be removed, and vice versa). This indicates that the model adopts a coarse-to-fine correction strategy to progressively remove the most significant artifacts. Then, we compare the correction performance between the model's adaptive order (inference) and the reverse order, as shown in Fig. 5 B. PSNR and SSIM across all bias field intensities confirm the superiority of the adaptive sequence.

4 Conclusion

In this paper, we propose ResMAP, a cascading framework for restoring MRIs with mixed artifacts. Our method addresses current challenges in MRI mixed artifact removal: the lack of paired mixed artifact data for training and the poor performance of models trained on individual artifact types when applied to mixed scenarios. To overcome these, ResMAP employs an iterative removal strategy guided by prompts retrieval. Experiments demonstrate superior performance of our method, and zero-shot experiments across varying field strengths further validate its robustness in diverse scenarios. Still, defining the clean standard for MRI remains further exploration for automatic termination of iterations.

Acknowledgments. This work was partially supported by AI4S Initiative and HPC Platform of ShanghaiTech University.

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

References

1. Callewaert, B., Jones, E.A., Himmelreich, U., Gsell, W.: Non-invasive evaluation of cerebral microvasculature using pre-clinical mri: principles, advantages and limitations. Diagnostics **11**(6), 926 (2021)
2. Chen, Z., Pawar, K., Ekanayake, M., Pain, C., Zhong, S., Egan, G.F.: Deep learning for image enhancement and correction in magnetic resonance imaging—state-of-the-art and challenges. Journal of Digital Imaging **36**(1), 204–230 (2023)
3. Conde, M.V., Geigle, G., Timofte, R.: Instructir: High-quality image restoration following human instructions. In: European Conference on Computer Vision. pp. 1–21. Springer (2024)
4. Deng, X., Dragotti, P.L.: Deep convolutional neural network for multi-modal image restoration and fusion. IEEE transactions on pattern analysis and machine intelligence **43**(10), 3333–3348 (2020)
5. Dietrich, O., Reiser, M.F., Schoenberg, S.O.: Artifacts in 3-t mri: physical background and reduction strategies. European journal of radiology **65**(1), 29–35 (2008)
6. Duffy, B.A., Zhang, W., Tang, H., Zhao, L., Law, M., Toga, A.W., Kim, H.: Retrospective correction of motion artifact affected structural mri images using deep learning of simulated motion. In: Medical imaging with deep learning (2018)
7. Duffy, B.A., Zhao, L., Sepehrband, F., Min, J., Wang, D.J., Shi, Y., Toga, A.W., Kim, H., Initiative, A.D.N., et al.: Retrospective motion artifact correction of structural mri images using deep learning improves the quality of cortical surface reconstructions. Neuroimage **230**, 117756 (2021)
8. Gibson, E., Li, W., Sudre, C., Fidon, L., Shakir, D.I., Wang, G., Eaton-Rosen, Z., Gray, R., Doel, T., Hu, Y., et al.: Niftynet: a deep-learning platform for medical imaging. Computer methods and programs in biomedicine **158**, 113–122 (2018)
9. Goldfryd, T., Gordon, S., Raviv, T.R.: Deep semi-supervised bias field correction of mr images. In: 2021 IEEE 18th international symposium on biomedical imaging (ISBI). pp. 1836–1840. IEEE (2021)

10. Jiang, J., Zuo, Z., Wu, G., Jiang, K., Liu, X.: A survey on all-in-one image restoration: Taxonomy, evaluation and future trends. arXiv preprint arXiv:2410.15067 (2024)
11. Jiang, N., Huang, Z., Sui, Y.: Explanation-driven cyclic learning for high-quality brain mri reconstruction from unknown degradation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 318–328. Springer (2024)
12. Kidoh, M., Shinoda, K., Kitajima, M., Isogawa, K., Nambu, M., Uetani, H., Morita, K., Nakaura, T., Tateishi, M., Yamashita, Y., et al.: Deep learning based noise reduction for brain mr imaging: tests on phantoms and healthy volunteers. Magnetic resonance in medical sciences **19**(3), 195–206 (2020)
13. Lei, Y., Li, J., Li, Z., Cao, Y., Shan, H.: Prompt learning in computer vision: a survey. Frontiers of Information Technology & Electronic Engineering **25**(1), 42–63 (2024)
14. Li, B., Liu, X., Hu, P., Wu, Z., Lv, J., Peng, X.: All-in-one image restoration for unknown corruption. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 17452–17462 (2022)
15. Lyu, M., Mei, L., Huang, S., Liu, S., Li, Y., Yang, K., Liu, Y., Dong, Y., Dong, L., Wu, E.X.: M4raw: A multi-contrast, multi-repetition, multi-channel mri k-space dataset for low-field mri research. Scientific Data **10**(1), 264 (2023)
16. Masoudnia, S., Ebrahimpour, R.: Mixture of experts: a literature survey. Artificial Intelligence Review **42**, 275–293 (2014)
17. Pérez-García, F., Sparks, R., Ourselin, S.: Torchio: a python library for efficient loading, preprocessing, augmentation and patch-based sampling of medical images in deep learning. Computer methods and programs in biomedicine **208**, 106236 (2021)
18. Potlapalli, V., Zamir, S.W., Khan, S.H., Shahbaz Khan, F.: Promptir: Prompting for all-in-one image restoration. Advances in Neural Information Processing Systems **36**, 71275–71293 (2023)
19. Smith, T.B.: Mri artifacts and correction strategies. Imaging in Medicine **2**(4), 445 (2010)
20. Sundaresan, V., Dinsdale, N.K.: Automated quality assessment using appearance-based simulations and hippocampus segmentation on low-field paediatric brain mr images. arXiv preprint arXiv:2410.06161 (2024)
21. Van Leemput, K., Maes, F., Vandermeulen, D., Suetens, P.: Automated model-based tissue classification of mr images of the brain. IEEE transactions on medical imaging **18**(10), 897–908 (1999)
22. Xiao, S., Liu, Z., Zhang, P., Muennighoff, N., Lian, D., Nie, J.Y.: C-pack: Packed resources for general chinese embeddings. In: Proceedings of the 47th international ACM SIGIR conference on research and development in information retrieval. pp. 641–649 (2024)
23. Yang, Z., Chen, H., Qian, Z., Yi, Y., Zhang, H., Zhao, D., Wei, B., Xu, Y.: All-in-one medical image restoration via task-adaptive routing. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 67–77. Springer (2024)
24. Zamir, S.W., Arora, A., Khan, S., Hayat, M., Khan, F.S., Yang, M.H.: Restormer: Efficient transformer for high-resolution image restoration. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 5728–5739 (2022)