

UniField: A Unified Field-Aware MRI Enhancement Framework

Yiyang Lin¹, Chenhui Wang², Zhihao Peng¹, and Yixuan Yuan¹

¹ Department of Electronic Engineering, the Chinese University of Hong Kong

² The Institute of Science and Technology for Brain-inspired Intelligence, Fudan University

yylin247@163.com, chenhuiwang21@m.fudan.edu.cn, yxyuan@ee.cuhk.edu.hk

Abstract. Magnetic Resonance Imaging (MRI) field-strength enhancement holds immense value for both clinical diagnostics and advanced research. However, existing methods typically focus on isolated enhancement tasks, such as specific 64mT-to-3T or 3T-to-7T transitions using limited subject cohorts, thereby failing to exploit the shared degradation patterns inherent across different field strengths and severely restricting model generalization. To address this challenge, we propose UniField, a unified framework integrating multiple modalities and enhancement tasks to mutually promote representation learning by exploiting these shared degradation characteristics. Specifically, our main contributions are threefold. Firstly, to overcome MRI data scarcity and capture continuous anatomical structures, UniField departs from conventional methods that treat 3D MRI volumes as independent 2D slices. Instead, we directly exploit comprehensive 3D volumetric information by leveraging pre-trained 3D foundation models, thereby embedding generalized and robust structural representations to significantly boost enhancement performance. In addition, to mitigate the spectral bias of mainstream flow-matching models that often over-smooth high-frequency details, we explicitly incorporate the physical mechanisms of magnetic fields to introduce a Field-Aware Spectral Rectification Mechanism (FASRM), tailoring customized spectral corrections to distinct field strengths. Finally, to resolve the fundamental data bottleneck, we organize and publicly release a comprehensive paired multi-field MRI dataset, which is an order of magnitude larger than existing datasets. Extensive experiments demonstrate our method’s superiority over state-of-the-art approaches, achieving an average improvement of approximately 1.81 dB in PSNR and 9.47% in SSIM. Codes and datasets are available at: <https://github.com/linyiyang98/UniField>.

Keywords: Unified MRI Field Strength Enhancement · Field-Aware Spectral Rectification Mechanism · Large-Scale Field-Enhanced Dataset.

* Yiyang Lin and Chenhui Wang contribute equally to this work.

** Corresponding author: Yixuan Yuan.

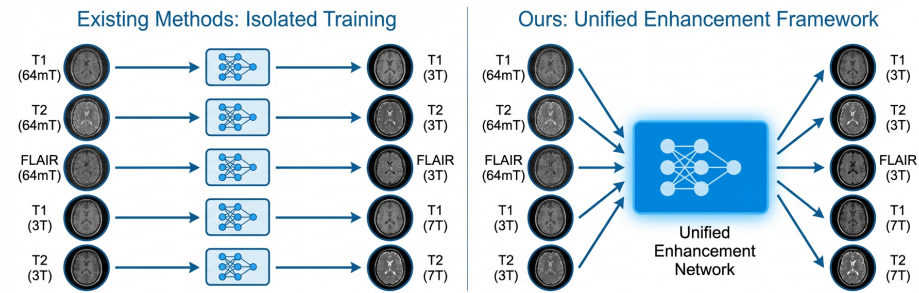


Fig. 1. Comparison of existing MRI enhancing method and our framework. (Left) Existing methods train isolated networks for specific modalities and tasks. (Right) Our method unifies all modalities and tasks within a single framework.

1 Introduction

Magnetic Resonance Imaging (MRI) [9,12,18] is crucial for medical diagnosis, particularly in brain imaging. In clinical practice, MRI utilizes various field strengths: ultra-low-field (64mT) scanners offer portability and accessibility, standard clinical-field (3T) provides high-resolution structural mapping, and ultra-high-field (7T) reveals unprecedented microscopic details. Computational field-strength enhancement techniques offer a transformative solution to bridge these hardware gaps. Upgrading low-cost 64mT scans to 3T quality democratizes high-fidelity diagnostics at the patient’s bedside, while enhancing clinical 3T images to 7T quality unlocks advanced neuroimaging capabilities without the exorbitant costs of ultra-high-field systems [7,11].

However, developing robust computational MRI enhancement models is severely bottlenecked by three main challenges. First, the fundamental data scarcity and isolated training paradigms remain major hurdles. Existing studies typically rely on merely a few dozen strictly paired cases [3,4,15], and predominantly train isolated models for specific modalities (*e.g.* T1, T2) or localized tasks (*e.g.* 64mT-to-3T). This fragmented approach inevitably causes overfitting (Fig. 1, left) and fails to exploit the shared degradation patterns inherent across different field strengths [6,19]. Second, conventional methods often treat 3D MRI volumes as independent 2D slices. This slice-by-slice processing discards crucial inter-slice continuous anatomical structures, severely limiting structural fidelity. Third, mainstream MRI enhancement methods, such as flow-matching based models, suffer from severe spectral bias. By overlooking the distinct physical mechanisms of magnetic fields at varying strengths, they tend to over-smooth high-frequency details that are vital for clinical diagnosis.

To address these intertwined challenges, we propose UniField, a unified framework that integrates multiple modalities and enhancement tasks to mutually promote representation learning. Specifically, our solutions are threefold, directly targeting the aforementioned bottlenecks. Firstly, to resolve the fundamental data scarcity and isolated training limitations, we organize and publicly

release a comprehensive paired multi-field MRI dataset [17,16], which is an order of magnitude larger than existing benchmarks, enabling our unified model to fully exploit shared degradation characteristics. In addition, to overcome the limitations of 2D processing, UniField departs from conventional slice-by-slice methods. Instead, we exploit comprehensive 3D volumetric information by leveraging pre-trained 3D foundation models, thereby embedding generalized and robust structural representations. Finally, to mitigate the spectral bias and preserve fine anatomical details, we explicitly incorporate the physical mechanisms of magnetic fields to introduce a Field-Aware Spectral Rectification Mechanism (FASRM), tailoring customized spectral corrections to distinct field strengths.

Our main contributions can be summarized as follows: (1) We propose UniField to unify diverse MRI modalities and enhancement tasks, leveraging shared intrinsic degradation patterns for mutual performance promotion. (2) To mitigate the spectral bias and over-smoothing inherent in traditional flow-matching, we introduce a novel dual-domain paradigm. It leverages the decoupled frequency domain to tailor customized spectral optimization schemes based on the distinct physical characteristics of varying field strengths. (3) We curate and release the largest paired multi-field MRI enhancement dataset to date. By precisely registering previously unaligned data, we increase benchmark volumes by an order of magnitude, establishing a robust foundation for future research.

2 Method

2.1 Overall Architecture of UniField

To overcome the data scarcity and poor generalization of isolated models, we propose UniField, a unified MRI field-strength enhancement framework (Fig. 2). This unified framework integrates diverse modalities (*e.g.* T1, T2, FLAIR) and cross-field enhancements—such as low-field (LF) 64mT to high-field (HF) 3T, and 3T to 7T—to mutually promote performance via shared degradation patterns. Furthermore, to embed robust structural priors, UniField operates within the latent space of FlashVSR [20], a pre-trained video super-resolution model.

During the forward process, flow-matching constructs a time-dependent latent state $\mathbf{z}_t$ at $t \in [0, 1]$ by interpolating between the clean HF target latent $\mathbf{z}_0 = \mathcal{E}(\mathbf{x}_{\text{HF}})$ and standard Gaussian noise $\mathbf{z}_1 \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. The core network $\mathcal{F}_\theta$ is trained to predict the velocity field driving this transition, explicitly conditioned on the current timestep t .

During inference, given an LF input $\mathbf{x}_{\text{LF}}$, initial noise $\mathbf{z}_1$, and a frozen UMT5-XXL [13]-encoded text condition $\mathbf{c}$ (formulated as “MRI {modality (*e.g.*, T1, T2, and FLAIR)} sequence enhancement from {input_field (*e.g.*, 64mT and 3T)} to {target_field (*e.g.*, 3T and 7T)} magnetic field”), we solve the empirical ODE:

$$\hat{\mathbf{x}}_{\text{HF}} = \mathcal{D}(\text{ODE}(\mathcal{F}_\theta(\mathbf{z}_1, \mathcal{E}(\mathbf{x}_{\text{LF}}), \mathbf{c}))), \quad (1)$$

where $\mathcal{E}(\cdot)$ and $\mathcal{D}(\cdot)$ are the frozen FlashVSR encoder and decoder. Specifically, the ODE solver $\text{ODE}(\cdot)$ numerically evaluates the time-dependent predicted velocity field $\mathcal{F}_\theta$ across timesteps from $t = 1$ to $t = 0$, yielding a clean latent

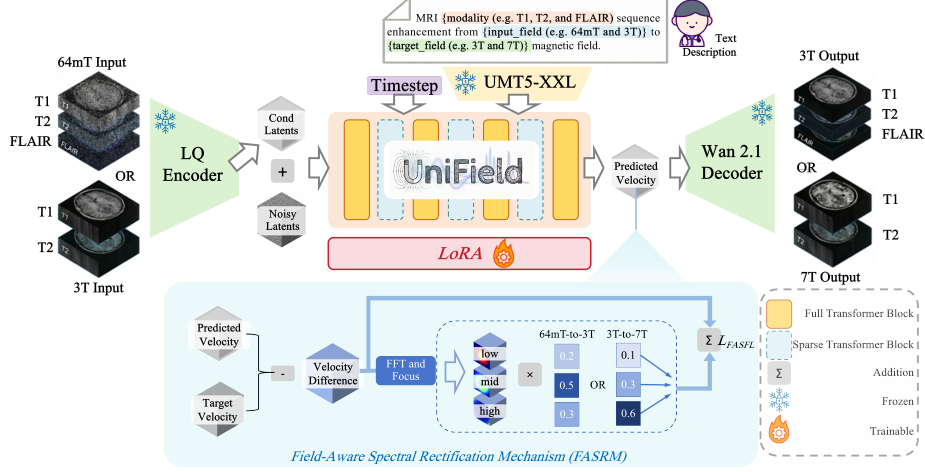


Fig. 2. Overview of our proposed UniField. Within the framework, the encoder and decoder utilize frozen pre-trained FlashVSR weights, while the central unified network is fine-tuned using LoRA.

representation $\mathbf{z}_0$ that $\mathcal{D}(\cdot)$ reconstructs into the target HF volume $\hat{\mathbf{x}}_{\text{HF}}$. To efficiently adapt $\mathcal{F}_\theta$ for 3D MRI, we employ Low-Rank Adaptation (LoRA) [14] alongside sparse attention [14], preventing long-range spatial dependencies from blurring critical local anatomy.

However, while this latent flow-matching process provides a powerful generative backbone, it inherently suffers from spectral bias, often resulting in over-smoothed structures. To mitigate this, we introduce a Field-Aware Spectral Rectification Mechanism (FASRM) (Fig. 2, blue box), which explicitly optimizes high-frequency details based on field-specific physical characteristics.

2.2 Unified Modeling via Video Super-Resolution Prior

Acquiring paired high-field MRI data is expensive and challenging. To overcome this data scarcity, we propose a dual approach: unified modeling to maximize data utilization, and an external prior to bypass training from scratch.

Unified Modeling. Lower-field MRI degradation patterns are highly consistent across modalities (e.g., T1, T2, FLAIR) and field transitions (e.g., 64mT-to-3T, 3T-to-7T). Consolidating these tasks into a single framework conditioned on modality and transition acts as implicit data augmentation. This enables the joint learning of shared enhancement features—such as noise suppression and detail recovery—drastically reducing the need for massive paired data per task.

Video Super-Resolution Prior. Training 3D generative models from scratch is heavily data-dependent. Because 3D MRI and 2D video share analogous spatiotemporal structures and enhancement goals, we introduce a pre-trained video prior to provide rich, generalized structural representations. Specif-

ically, we initialize our framework with FlashVSR [20]. To efficiently adapt this powerful generative prior to 3D MRI, we freeze the encoder and decoder, and employ Low-Rank Adaptation (LoRA) to fine-tune the main body.

2.3 Field-Aware Spectral Rectification Mechanism (FASRM)

To overcome the spectral bias and over-smoothing in traditional flow-matching, we propose a field-aware spatial-frequency loss (FASFL) within a field-aware spectral rectification mechanism (FASRM). Unlike uniform or task-agnostic spectral penalties, FASFL dynamically modulates frequency-domain optimization based on the physical properties of the source and target magnetic fields.

Specifically, FASRM introduces a field-conditioned spectral rectification scheme in the frequency domain to address the fundamentally different physical constraints of diverse enhancement tasks. For example, in extreme low-to-high field transitions (e.g., 64mT-to-3T), high-frequency features of the target high-field images lack corresponding structural hints in the low-field source. Over-penalizing the high-frequency band would force the model to blindly hallucinate these structures. Thus, FASRM automatically relaxes high-frequency constraints to prioritize structural fidelity. As another example, in high-to-ultra-high field transitions (e.g., 3T-to-7T), target images frequently suffer from inherent physical artifacts, such as B1 inhomogeneity, which predominantly manifest in the low-frequency spectrum. To prevent the model from memorizing and reproducing these artifacts, FASRM suppresses the low-frequency learning weights. By explicitly embedding these field-specific physical priors, our method effectively mitigates both hallucinations and artifact propagation.

In summary, the unified objective, which simultaneously enforces spatial fidelity and spectral rectification, is defined as:

$$\begin{aligned} \mathcal{L}_{FASFL} = & \|v_\theta - v\|_1 \\ & + \lambda_{freq} \sum_{b \in \mathcal{B}} \frac{w_b}{|M_b|} \|M_b \odot (|F(v_\theta) - F(v)|^\alpha \odot |F(v_\theta) - F(v)|^2)\|_1, \end{aligned} \quad (2)$$

where v_θ and v are the predicted and target velocity fields, and $F(\cdot)$ denotes the Fast Fourier Transform (FFT). $\mathcal{B} = \{low, mid, high\}$ denotes the partitioned frequency bands with binary masks M_b , where $|M_b|$ represents the number of elements in the mask. Crucially, the field-conditioned weights w_b are designed for adapting to the specific field transition, α modulates the degree of focus on challenging spectra, and λ_{freq} controls the strength of spectral rectification.

3 Experiments

3.1 Dataset and Experimental Details

To overcome the limitations of small-scale, single-center studies, we curate a comprehensive multi-center, multi-task MRI dataset sourced from five institutions.

Table 1. Dataset overview. For unregistered MRI pairs, we perform registration.

Task	Dataset	Modalities	Registered	Total Train Test		
64mT-to-3T	ULF-EnC [8]	T1, T2, FLAIR	Yes	50	40	10
	Leiden [2]	T1, T2, FLAIR	No	10	8	2
	KCL [15]	T1, T2	No	23	18	5
3T-to-7T	UNC [3]	T1, T2	Yes	10	8	2
	BNU [4]	T1, T2	No	20	16	4

Table 2. Quantitative comparison of our UniField and other competing methods on the 64mT-to-3T task for T1, T2, and FLAIR modalities. NRMSE ($\downarrow$) and SSIM ($\uparrow$) are in %, PSNR ($\uparrow$) in dB, and LPIPS ($\downarrow$) is unitless ($\uparrow/\downarrow$: higher/lower is better). Best and second-best results are **bolded** and underlined, respectively.

Methods	T1				T2				FLAIR			
	NRMSE	PSNR	SSIM	LPIPS	NRMSE	PSNR	SSIM	LPIPS	NRMSE	PSNR	SSIM	LPIPS
MO-U-NET(2025)	11.4	18.98	52.9	0.28	<u>9.5</u>	<u>20.58</u>	59.2	0.23	<u>10.9</u>	<u>19.32</u>	53.3	0.25
MSFA(2024)	<u>11.3</u>	<u>19.05</u>	<u>67.9</u>	<u>0.26</u>	9.9	20.27	<u>71.6</u>	0.22	<u>10.9</u>	19.31	<u>65.6</u>	<u>0.24</u>
LowGAN(2025)	12.8	18.03	66.6	0.19	11.4	18.98	65.8	<u>0.20</u>	14.1	17.02	63.2	0.20
FlashVSR(2026)	12.1	18.41	63.4	0.30	12.0	18.54	66.3	0.23	15.1	16.61	60.6	0.30
UniField (Ours)	10.5	19.75	72.8	0.19	8.4	21.57	74.9	0.16	9.5	20.50	71.5	0.20

Featuring rigorous preprocessing and precise cross-field registration, this dataset enables the systematic evaluation of both 64mT $\rightarrow$ 3T and 3T $\rightarrow$ 7T enhancements. As detailed in Table 1, all experiments utilize a uniform 8:2 train-test split across each center. Preprocessing includes MONAI intensity normalization (0.5th–99.5th percentiles), 1mm z-resampling, and $256 \times 256 \times 160$ resizing.

We set $\lambda_{freq} = 0.1$, $\alpha = 1.0$, and frequency weights w_b to $[0.2, 0.5, 0.3]$ (64mT-to-3T) and $[0.1, 0.3, 0.6]$ (3T-to-7T). Implemented in PyTorch, the model was trained on an RTX A6000 (batch size 1, 1,000 iterations) using Adam ($\beta_1 = 0.5$, $\beta_2 = 0.999$, initial $lr = 10^{-4}$ with linear decay after 500).

3.2 Comparative Experiments

In our comparative experiments, we evaluate our method against several baselines, including the U-Net-based MO-U-NET [1], GAN-based MSFA [5] and LowGAN [10], and diffusion-based FlashVSR [20]. As observed in Fig. 3, MO-U-NET and MSFA generate highly blurred images, particularly in the red boxes of the 64mT-to-3T task. Meanwhile, LowGAN mistakenly generates normal brain tissue as a bright white hyperintense artifact, as shown in the red circle of the T2 3T-to-7T task. In addition, although FlashVSR produces better results than the previous methods, its error map still exhibits darker colors, especially in high-frequency details. In contrast, our method successfully addresses most of these issues, recovering sharp tissue boundaries and rich textures that are highly consistent with the ground truth, which is further validated by having the lowest residual values in the error maps. Quantitatively, as shown in Tables 2 and 3, our method achieves the best performance across the vast majority of evaluation metrics. This comprehensive superiority in both visual fidelity and specific numerical gains fully demonstrates the effectiveness of our proposed approach.

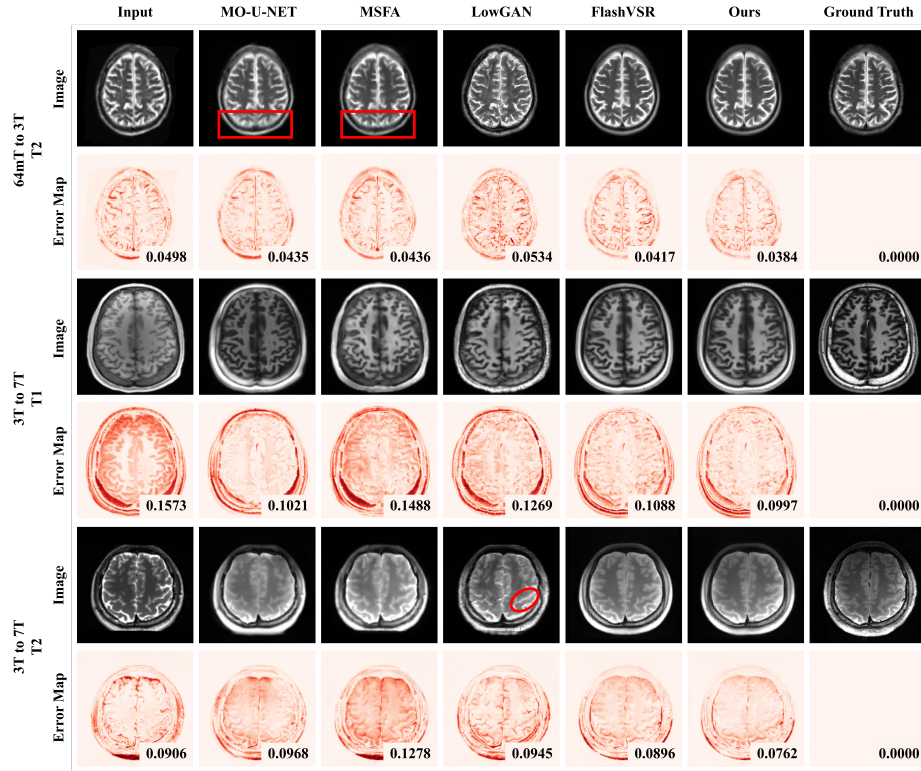


Fig. 3. Visual comparison with competing methods. Error maps show the absolute difference from the ground truth, where darker colors indicate larger errors (mean pixel error at bottom right). Red annotations highlight defects in competing methods.

Table 3. Quantitative comparison of our UniField and other competing methods on the 3T-to-7T task for T1 and T2 modalities. NRMSE ($\downarrow$) and SSIM ($\uparrow$) are in %, PSNR ($\uparrow$) in dB, and LPIPS ($\downarrow$) is unitless ($\uparrow/\downarrow$: higher/lower is better). Best and second-best results are **bolded** and underlined, respectively.

Methods	T1				T2			
	NRMSE	PSNR	SSIM	LPIPS	NRMSE	PSNR	SSIM	LPIPS
MO-U-NET (2025)	11.9	18.51	60.8	0.28	8.6	21.32	59.7	0.27
MSFA (2024)	<u>10.5</u>	<u>19.66</u>	66.8	<u>0.23</u>	<u>8.1</u>	<u>21.82</u>	<u>65.0</u>	<u>0.24</u>
LowGAN (2025)	15.4	16.26	54.0	0.25	10.1	19.97	56.0	0.25
FlashVSR (2026)	11.8	18.63	63.0	<u>0.23</u>	9.5	20.49	58.6	0.24
UniField (Ours)	9.9	20.12	<u>66.1</u>	0.21	7.5	22.55	72.1	0.18

3.3 Ablation Experiments

We first compare a unified multi-modality model against modality-specific models. As shown in Table 4, the unified model consistently outperforms individual models across 64mT-to-3T and 3T-to-7T. Since lower-field degradation patterns are similar across modalities, joint training acts as data augmentation. Additionally, a unified framework reduces storage and streamlines clinical workflows.

Table 4. Quantitative ablation on multi-modality unification. NRMSE ($\downarrow$) and SSIM ($\uparrow$) are in %, PSNR ($\uparrow$) in dB, and LPIPS ($\downarrow$) is unitless ($\uparrow/\downarrow$: higher/lower is better). Best results are **bolded**.

Methods	T1				T2				FLAIR			
	NRMSE	PSNR	SSIM	LPIPS	NRMSE	PSNR	SSIM	LPIPS	NRMSE	PSNR	SSIM	LPIPS
<i>Task: 64mT-to-3T</i>												
T1 Only	12.1	18.41	63.4	0.30	-	-	-	-	-	-	-	-
T2 Only	-	-	-	-	12.0	18.54	66.3	0.23	-	-	-	-
FLAIR Only	-	-	-	-	-	-	-	-	15.1	16.61	60.6	0.30
Unified modalities	11.7	18.71	65.1	0.28	9.4	20.58	68.6	0.23	11.1	19.17	64.5	0.30
+ Unified tasks	10.7	19.06	72.1	0.20	8.9	21.05	72.8	0.18	11.9	19.06	66.7	0.24
+ FASRM (Ours)	10.5	19.75	72.8	0.19	8.4	21.57	74.9	0.16	9.5	20.50	71.5	0.20
<i>Task: 3T-to-7T</i>												
T1 Only	11.8	18.63	63.0	0.23	-	-	-	-	-	-	-	-
T2 Only	-	-	-	-	9.5	20.49	58.6	0.24	-	-	-	-
Unified modalities	11.0	19.25	64.4	0.21	9.0	20.97	61.1	0.21	-	-	-	-
+ Unified tasks	10.6	19.53	65.8	0.21	9.2	20.88	66.6	0.19	-	-	-	-
+ FASRM (Ours)	9.9	20.12	66.1	0.21	7.5	22.55	72.1	0.18	-	-	-	-

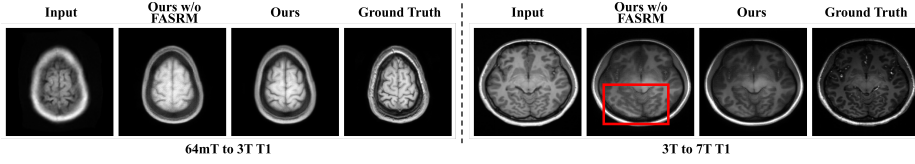


Fig. 4. Visualization of ablation experiments on FASRM. Results for both 64mT-to-3T and 3T-to-7T tasks demonstrate the module’s effectiveness. Red boxes highlight severe defects in the baselines without FASRM.

Building upon this, a single unified cross-task model significantly surpasses task-specific networks (Table 4). Since field enhancement inherently involves analogous mappings—such as noise suppression and high-frequency recovery—joint training enables the network to learn a robust, generalized prior, effectively transferring structural representations across different field-strength domains.

Finally, we validate the proposed FASRM. As illustrated in Fig. 4, omitting FASRM leads to poor high-frequency reconstruction, yielding blurred images in the 64mT-to-3T task and anatomically inconsistent structures in the 3T-to-7T task. Incorporating FASRM effectively resolves these issues, producing sharper, anatomically faithful results. The quantitative metrics in Table 4 further corroborate the performance improvements brought by FASRM.

4 Conclusion

In this paper, we introduce UniField, a unified framework that advances MRI field-strength enhancement by overcoming data scarcity and generalization bottlenecks. We construct the largest registered multi-field dataset to date and leverage video super-resolution priors with shared degradation patterns to maximize data efficiency. Furthermore, our novel field-aware spectral rectification mechanism resolves the spectral bias inherent in generative models, restoring sharp high-frequency anatomical details. Future work will expand this paradigm

to diverse anatomical organs, magnetic field strengths, and disease types, paving the way for universally accessible, high-quality clinical diagnostics.

Acknowledgement. This work was supported by NSFC/RGC Joint Research Scheme N_CUHK4126/25, Innovation and Technology Fund Mainland-Hong Kong Joint Funding Scheme MHP/173/24.

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

References

1. Baljer, L., Zhang, Y., Bourke, N.J., Donald, K.A., Bradford, L.E., Ringshaw, J.E., Williams, S.R., Deoni, S.C., Williams, S.C., Team, K.S.S., et al.: Ultra-low-field paediatric mri in low-and middle-income countries: Super-resolution using a multi-orientation u-net. *Human Brain Mapping* **46**(1), e70112 (2025)
2. van den Broek, R.B., Lena, B., Webb, A.: Paired 64mt and 3t brain mri scans of healthy subjects for neuroimaging research. Preprint (2025)
3. Chen, X., Qu, L., Xie, Y., Ahmad, S., Yap, P.T.: A paired dataset of t1-and t2-weighted mri at 3 tesla and 7 tesla. *Scientific Data* **10**(489) (2023)
4. Chu, L., Ma, B., Dong, X., He, Y., Che, T., Zeng, D., Zhang, Z., Li, S.: A paired dataset of multi-modal mri at 3 tesla and 7 tesla with manual hippocampal subfield segmentations. *Scientific Data* **12**(260) (2025)
5. Cui, Q., Tosun, D., Mukherjee, P., Abbasi-Asl, R.: 7t mri synthesization from 3t acquisitions. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 35–44. Springer (2024)
6. Fei, B., Lyu, Z., Pan, L., Zhang, J., Yang, W., Luo, T., Zhang, B., Dai, B.: Generative diffusion prior for unified image restoration and enhancement. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. pp. 9935–9946 (2023)
7. He, Z., Li, W., Jiang, Y., Peng, Z., Wang, P., Li, X., Liu, T., Han, J., Zhang, T., Yuan, Y.: F2tnet: Fmri to t1w mri knowledge transfer network for brain multi-phenotype prediction. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 265–275. Springer (2024)
8. Islam, K.T., Zhong, S., Zakavi, P., et al.: Improving portable low-field mri image quality through image-to-image translation using paired low-and high-field images. *Scientific Reports* **13**(1), 21183 (2023)
9. Katti, G., Ara, S.A., Shireen, A.: Magnetic resonance imaging (mri)—a review. *Int. J. Dent. Clin* **3**(1), 65–70 (2011)
10. Lucas, A., Arnold, T.C., Okar, S.V., Vadali, C., Kawatra, K.D., Ren, Z., Cao, Q., Shinohara, R.T., Schindler, M.K., Davis, K.A., et al.: Multisequence 3-t image synthesis from 64-mt low-field-strength mri using generative adversarial networks in multiple sclerosis. *Radiology* **315**(1), e233529 (2025)
11. Peng, Z., He, Z., Jiang, Y., Wang, P., Yuan, Y.: Gbt: Geometric-oriented brain transformer for autism diagnosis. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 142–152. Springer (2024)
12. Peng, Z., Wang, C., Liu, S., Liang, Z., Ye, Z., Ju, M., Woo, P., Yuan, Y.: Omnibrainbench: A comprehensive multimodal benchmark for brain imaging analysis across multi-stage clinical tasks. arXiv preprint arXiv:2511.00846 (2025)

13. Pirhadi, M.J., Mirzaei, M., Eetemadi, S.: Cvt5: Using compressed video encoder and umt5 for dense video captioning. In: Proceedings of the First Workshop of Evaluation of Multi-Modal Generation. pp. 10–23 (2025)
14. Sun, Z., Yang, H., Liu, K., Yin, Z., Li, Z., Xu, W.: Recent advances in lora: A comprehensive survey. *ACM Transactions on Sensor Networks* **18**(4), 1–44 (2022)
15. Váša, F., Bernallick, C., Bourke, N.J., Padormo, F., Baljer, L., Briski, U., Cawley, P., Arichi, T., Wood, T.C., Lythgoe, D.J., et al.: Ultra-low-field brain mri morphometry: Test–retest reliability and correspondence to high-field mri. *Imaging Neuroscience* **3** (2025)
16. Wang, C., Lei, Y., Chen, T., Zhang, J., Li, Y., Shan, H.: Hope: Hybrid-granularity ordinal prototype learning for progression prediction of mild cognitive impairment. *IEEE Journal of Biomedical and Health Informatics* **28**(11), 6429–6440 (2024)
17. Wang, C., Piao, S., Huang, Z., Gao, Q., Zhang, J., Li, Y., Shan, H., Initiative, A.D.N., et al.: Joint learning framework of cross-modal synthesis and diagnosis for alzheimer’s disease by mining underlying shared modality information. *Medical Image Analysis* **91**, 103032 (2024)
18. Wang, C., Zheng, B., Bao, L., Peng, Z., Woo, P.Y., Shan, H., Yuan, Y.: Brain-wm: Brain glioblastoma world model. *arXiv preprint arXiv:2603.07562* (2026)
19. Yang, K., Hu, T., Dai, K., Chen, G., Cao, Y., Dong, W., Wu, P., Zhang, Y., Yan, Q.: Crnet: A detail-preserving network for unified image restoration and enhancement task. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 6086–6096 (2024)
20. Zhuang, J., Guo, S., Cai, X., Li, X., Liu, Y., Yuan, C., Xue, T.: Flashvsr: Towards real-time diffusion-based streaming video super-resolution. *arXiv preprint arXiv:2510.12747* (2025)