PiMoE: Physics-informed Mixture-of-Experts for Accelerated MRI Reconstruction

Kyuri Kim[1], Yoonho Na[1], and Sung-Joon Ye[1,2]

[1] Department of Applied Bioengineering, Seoul National University, Seoul 08826, Republic of Korea
[2] Research Institute for Convergence Science, Seoul National University, Seoul 08826, Republic of Korea
{kyurikim, yoonho94.na, sye}@snu.ac.kr

Abstract. Deep unrolling networks have established a strong foundation for accelerated MRI reconstruction, yet they allocate uniform network capacity across all spatial regions, disregarding the inherently heterogeneous nature of reconstruction difficulty. In practice, this heterogeneity is governed by two distinct sources: physics-induced factors such as spatially varying aliasing and encoding instability determined by the acquisition operator, and anatomy-driven factors arising from structural complexity. We propose Physics-informed Mixture-of-Experts (PiMoE), a framework that replaces the monolithic learned prior with a set of specialized experts whose routing is guided by acquisition physics. Specifically, we derive three physical descriptors, aliasing susceptibility, phase alignment, and magnitude imbalance, directly from the undersampling mask and coil sensitivities without requiring ground-truth images. These descriptors are unified into a difficulty score that adaptively modulates the coupling between physics-based and feature-based routing. As local ill-posedness intensifies, routing authority shifts toward the physics-informed pathway, while feature-based routing dominates in well conditioned regions. Experiments on the fastMRI brain and CMRxRecon cardiac datasets demonstrate that PiMoE achieves consistent improvements over state-of-the-art methods across 4× and 8× acceleration, with generalization to unseen modalities and views.

Keywords: MRI Reconstruction · Mixture of Experts · Physics-informed Routing · Deep Unfolding Networks

1 Introduction

Magnetic resonance imaging (MRI) provides superior soft-tissue contrast and non-invasiveness, yet the physical constraints of sequential k-space acquisition inherently limit scan speed [2]. Undersampling techniques mitigate this by partially acquiring data, and recent deep learning-based unrolling methods [1,3,17,20] have established a performance standard by combining data consistency operations with learned priors. However, these approaches do not explicitly account for the spatially inhomogeneous nature of reconstruction difficulty, allocating uniform network capacity across all regions. In practice, even under identical acceleration factors, certain regions exhibit persistent aliasing or encoding insufficiency while others are restored reliably [10]. This spatial inhomogeneity reflects the combined influence of local acquisition conditions and anatomical complexity, suggesting that MRI reconstruction is fundamentally a *heterogeneous inverse problem* requiring spatially varying strategies.

Specifically, *physical difficulty* is determined by acquisition conditions: the global interference structure dictated by the sampling pattern and the local encoding stability governed by coil geometry define the conditioning of the inverse problem independently of image content, and are deterministically fixed at the time of acquisition. *Anatomical difficulty*, in contrast, is driven by the structural complexity of the image. The diversity of local features modulates the strength of regularization required for faithful restoration. A successful reconstruction framework must therefore explicitly reflect both the physical acquisition structure and anatomical characteristics. While recent studies have proposed adaptive designs leveraging uncertainty [25] or attention mechanisms [4,6], these approaches rely entirely on latent features, rendering them vulnerable to distribution shift. Moreover, as cascades deepen, the accumulated feature bias can destabilize the adaptation strategy. Fundamentally, such data-driven representations prioritize *what is observed* while overlooking the physical reality of *how the data was acquired*.

The Mixture-of-Experts (MoE) paradigm [16] offers a structurally well-suited solution to this spatial heterogeneity. By selectively activating only a subset of experts per input, MoE promotes functional specialization of individual experts [5,8,13,21], providing a natural framework for distributing a heterogeneous inverse problem across a discrete set of specialists. In MRI reconstruction specifically, the spatially varying ill-posedness gives rise to distinct degradation regimes, regions dominated by severe aliasing versus those limited by encoding insufficiency, making discrete expert partitioning better suited than the continuous adaptation of a single network. However, unlike general computer vision tasks [12,13,22], the reconstruction difficulty for the same anatomical structure can vary drastically depending on the acquisition environment. The ill-posedness determined by undersampling patterns and coil sensitivities is distributed heterogeneously across the image, creating reconstruction bottlenecks that are difficult to capture from image features alone. Consequently, routing that relies solely on data-driven features without explicit physical grounding leads to suboptimal expert assignment.

To overcome these limitations, we propose the **Physics-informed Mixture-of-Experts (PiMoE)** framework. PiMoE adopts a dual routing architecture that independently processes latent image features and physics-based descriptors derived deterministically from the acquisition conditions. A unified difficulty score adaptively modulates the authority allocation between the two routers. As the local ill-posedness intensifies, routing authority is progressively transferred to the physics-informed pathway, ensuring robust expert assignment even under severe degradation conditions. Our contributions are summarized as follows:

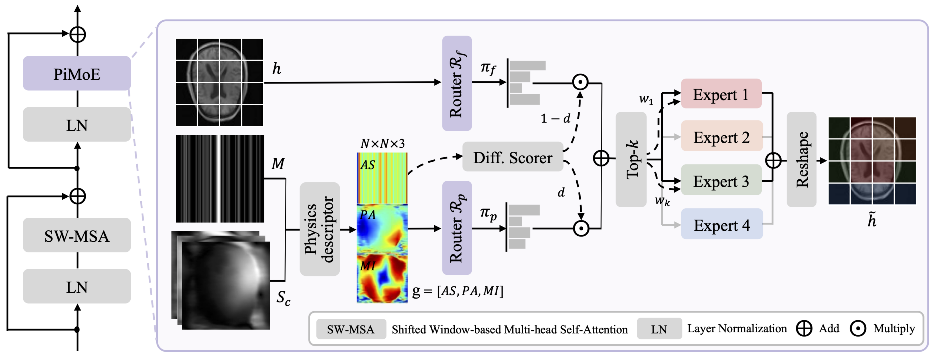

Fig. 1. Overview of PiMoE. A unified difficulty score d derived from physical descriptors, adaptively shifts routing authority between the feature router $\mathcal{R}_f$ and physics router $\mathcal{R}_p$, enabling token-wise expert specialization tailored to the local degradation profile (illustrated with top-k=2).

1. We propose PiMoE, a spatially adaptive MRI reconstruction framework that integrates three acquisition-derived physical descriptors into a dual-routing expert architecture. (Sec. 2.2)
2. We introduce a difficulty-aware fusion mechanism that dynamically modulates routing authority between data-driven and physics-driven pathways based on a unified ill-posedness score. (Sec. 2.3)
3. Experiments on the two open datasets demonstrate consistent improvements over existing methods across multiple acceleration settings, with robust generalization to unseen modalities and views. (Sec. 3)

2 Proposed Method

2.1 Problem Formulation

Multi-coil MRI reconstruction seeks to recover an image $x \in \mathbb{C}^n$ from undersampled k-space measurements $\tilde{k} = \{y_c\}_{c=1}^{C}$, where $y_c = M\mathcal{F}S_c x + \epsilon_c$ with undersampling mask M, Fourier operator $\mathcal{F}$, and coil sensitivities $\{S_c\}$. Denoting the forward encoding operator $A : x \mapsto \{M\mathcal{F}S_c x\}_{c=1}^{C}$, reconstruction is posed as $\hat{x} = \arg\min_x \|A(x) - \tilde{k}\|_2^2 + \mathcal{R}(x)$. We adopt a deep unrolling formulation that alternates a data-consistency (DC) update and a learned prior for T iterations:

$$z^t = \mathcal{D}^{(t)}(x^t; \tilde{k}), \qquad x^{t+1} = z^t + f_\phi^{(t)}(z^t), \tag{1}$$

where $\mathcal{D}^{(t)}$ enforces measurement consistency and $f_\phi^{(t)}$ is a learned regularization prior. In this formulation, acquisition physics enters exclusively through the data consistency operator, while the prior remains entirely data-driven. However, a single monolithic prior must simultaneously accommodate regions with vastly different degradation characteristics within a shared parameter space, limiting its capacity to specialize. To address this, we replace $f_\phi^{(t)}$ with a Physics-informed Mixture-of-Experts (PiMoE):

$$x^{t+1} = z^t + \sum_{e=1}^{E} \pi_e^t \cdot f_{\phi_e}^{(t)}(z^t), \tag{2}$$

where π_e^t denotes the gating weights that dynamically allocate E specialized experts according to the local conditioning of A, enabling each expert to develop restoration strategies tailored to distinct degradation profiles.

2.2 Acquisition-based Ill-posedness Descriptors

To characterize the spatially varying ill-posedness, we derive three descriptors directly from the acquisition parameters M and estimated coil sensitivities $\{S_c\}_{c=1}^C$. These descriptors decompose reconstruction difficulty into two complementary aspects: global interference dictated by sampling trajectories and local conditioning arising from receiver coil geometry.

Global Interference. Undersampling induces aliasing by superimposing signals through the point spread function (PSF) sidelobes [9,18]. We quantify this global interference at location $\mathbf{r}$ by defining **Aliasing Susceptibility (AS)**:

$$\mathrm{AS}(\mathbf{r}) = (\mathcal{P}_{\text{side}} \circledast \mathrm{RSS})(\mathbf{r}), \quad \text{where} \quad \mathcal{P}_{\text{side}} = |\mathcal{F}^{-1}\{M\}| \cdot \Omega_\tau(\mathbf{r}). \tag{3}$$

Here, $\circledast$ denotes the convolution operator, and $\Omega_\tau(\mathbf{r})$ is a mask that nulls the central mainlobe region $|\mathbf{r}| \leq \tau$. $\mathrm{RSS}(\mathbf{r}) = \sqrt{\sum_c |S_c(\mathbf{r})|^2}$ represents the root-sum-of-squares coil combination. Physically, the convolution aggregates aliasing contributions from all source locations $\mathbf{r}'$: $\mathcal{P}_{\text{side}}(\mathbf{r}-\mathbf{r}')$ quantifies how much signal at $\mathbf{r}'$ is folded onto $\mathbf{r}$ through the PSF sidelobes, while $\mathrm{RSS}(\mathbf{r}')$ serves as a content-agnostic proxy for local signal strength derived solely from coil sensitivities. The resulting $\mathrm{AS}(\mathbf{r})$ thus estimates the total aliasing energy incident on each spatial location without requiring access to the ground-truth image.

Local Conditioning. The stability of the inverse problem further depends on the linear independence of the coil sensitivity vectors [7]. The g-factor [11,14] consolidates all encoding conditions into a single scalar measure of noise amplification, but this aggregation discards the structural information needed for discriminative routing. We instead characterize local conditioning through two descriptors that preserve distinct aspects of the encoding geometry, providing the router with richer input than a single collapsed measure. First, **Phase Alignment (PA)** captures the degree of phase redundancy among coils. When coil phases are nearly identical, the encoding matrix becomes rank-deficient, exacerbating noise amplification. We quantify this as:

$$\mathrm{PA}(\mathbf{r}) = \left| \frac{1}{C} \sum_{c=1}^{C} e^{i\angle S_c(\mathbf{r})} \right|, \tag{4}$$

where values approaching unity indicate poor phase diversity. Complementarily, **Magnitude Imbalance (MI)** quantifies the distribution of fractional coil power $p_c(\mathbf{r})$, as conditioning also suffers when the signal is dominated by a limited subset of coils:

$$\mathrm{MI}(\mathbf{r}) = 1 - \left(C \sum_{c=1}^{C} p_c(\mathbf{r})^2 \right)^{-1}, \quad \text{where} \quad p_c(\mathbf{r}) = \frac{|S_c(\mathbf{r})|^2}{\sum_{c'} |S_{c'}(\mathbf{r})|^2}. \tag{5}$$

Here, $\mathrm{MI} = 0$ corresponds to balanced coil contributions. MI increases monotonically as a single coil dominates more strongly, saturating at $(C-1)/C$, which approaches 1 for large C. Together, these three descriptors provide spatially resolved maps of distinct degradation sources, aliasing interference (AS), phase redundancy (PA), and encoding insufficiency (MI), which serve as the physical basis for expert routing.

2.3 Physics-informed Mixture-of-Experts (PiMoE)

Difficulty Scorer. The three descriptors are aggregated into a unified difficulty score $d^{(t)} \in [0, 1]$, representing the probability that a given spatial token lies in a physically challenging region. This score modulates the balance between data-driven and physics-informed routing. Since the physical descriptors $\mathbf{g}$ remain fixed across cascades, the cascade-wise evolution of routing authority is governed by per-cascade learnable parameters $\mathbf{w}^{(t)} \in \mathbb{R}^3$ and $b^{(t)} \in \mathbb{R}$:

$$d^{(t)} = \sigma\big(\mathbf{g}^\top \operatorname{softplus}(\mathbf{w}^{(t)}) + b^{(t)}\big). \tag{6}$$

By applying the softplus function to $\mathbf{w}^{(t)}$, we enforce a monotonicity constraint ($w_i \geq 0$), ensuring that $d^{(t)}$ strictly increases as aliasing interference, phase redundancy, or encoding insufficiency worsens. This allows the network to learn distinct difficulty thresholds at each unrolling stage.

Dual Routing Mechanism. The PiMoE framework employs a dual-routing scheme to integrate anatomical context with physical grounding. Two independent routers, a feature router $\mathcal{R}_f$ and a physics router $\mathcal{R}_p$, process their respective inputs via learnable linear projections:

$$\pi_f = \mathcal{R}_f(\mathbf{h}) = \mathrm{Softmax}(W_h \mathbf{h}), \qquad \pi_p = \mathcal{R}_p(\mathbf{g}) = \mathrm{Softmax}(W_g \mathbf{g}), \tag{7}$$

where $\mathbf{h} \in \mathbb{R}^D$ is a latent feature vector and $\mathbf{g} \in \mathbb{R}^3$ is the physical descriptor. Crucially, while d compresses the three descriptors into a single scalar for authority modulation, the physics router $\mathcal{R}_p$ receives the full descriptor vector $\mathbf{g}$, preserving the fine-grained spatial information necessary for discriminative expert selection. The final integrated routing distribution π is obtained by interpolating between the two routing opinions:

$$\pi = (1-d)\pi_f + d\pi_p. \tag{8}$$

Intuitively, d shifts routing authority from data-driven cues toward physics-informed priors as the local inverse problem becomes increasingly ill-posed. Since

Table 1. Quantitative comparison on the fastMRI and CMRxRecon datasets with different acceleration factors in terms of PSNR, SSIM, and NMSE.

		4× Acceleration			8× Acceleration		
Dataset	Method	PSNR↑	SSIM↑	NMSE↓	PSNR↑	SSIM↑	NMSE↓
fastMRI	Zero-filled	30.57	0.7755	0.0518	27.19	0.6695	0.1071
	U-Net [15]	37.22	0.9267	0.0103	34.15	0.8965	0.0208
	E2E-VarNet [17]	38.70	0.9230	0.0083	34.70	0.8876	0.0193
	HUMUSNet [1]	40.12	0.9371	0.0046	36.87	0.9172	0.0090
	Ours	**40.25**	**0.9421**	**0.0044**	**37.07**	**0.9228**	**0.0088**
CMRxRecon	Zero-filled	29.73	0.6447	0.1868	26.75	0.5688	0.3595
	U-Net [15]	36.41	0.9272	0.0479	33.05	0.8653	0.1161
	E2E-VarNet [17]	40.83	0.9316	0.0234	36.17	0.8914	0.0469
	HUMUSNet [1]	41.92	0.9462	0.0129	36.29	0.9028	0.0446
	Ours	**42.44**	**0.9484**	**0.0110**	**37.73**	**0.9172**	**0.0339**

the routing module operates within a deeply non-linear unrolling cascade, the linear interpolation preserves the interpretability of d as a physically grounded modulator, while the subsequent top-k selection provides the necessary sparsification for sharp expert assignment.

3 Experiments

Dataset. Experiments were conducted on the fastMRI brain [23] and CMRxRecon cardiac [19] datasets. The fastMRI set comprises 1,578 volumes (FLAIR, T1, T1Pre, T2; 1,192/194/192 train/valid/test split) center-cropped to 320×320 with 2–24 coils, introducing substantial variability in encoding conditions. The CMRxRecon set comprises 1,591 volumes (T1, T2, cine in SAX/LAX/LVOT; 953/298/340 split) center-cropped to 256×256 with 10 virtual coils [24]. For both datasets, retrospective random Cartesian undersampling was applied at 4× and 8× acceleration, using 8% and 4% center lines as the autocalibration signal.

Implementation Details. All models were implemented in PyTorch on NVIDIA RTX 3090 GPUs. The unrolled network employs a HUMUSNet backbone with $T{=}8$ cascades, optimized with SSIM loss using Adam at a learning rate of $1{\times}10^{-4}$ until validation convergence. Following the encoder-decoder structure, PiMoE replaces the six feed-forward layers in the decoder's three Transformer blocks with $E{=}8$ experts and top-2 routing. Aliasing susceptibility was computed along the phase-encoding axis with the central mainlobe nulled.

Comparison with State-of-the-Art. Table 1 compares PiMoE against U-Net [15], E2E-VarNet [17], and HUMUSNet [1] under identical single-slice settings. At 8× acceleration, PiMoE improves PSNR over the HUMUSNet baseline from 36.87 dB to 37.07 dB on fastMRI and from 36.29 dB to 37.73 dB on CMRxRecon, with consistent gains in SSIM and NMSE across both datasets. This trend extends to the 4× setting as well, though the improvement is more pro-

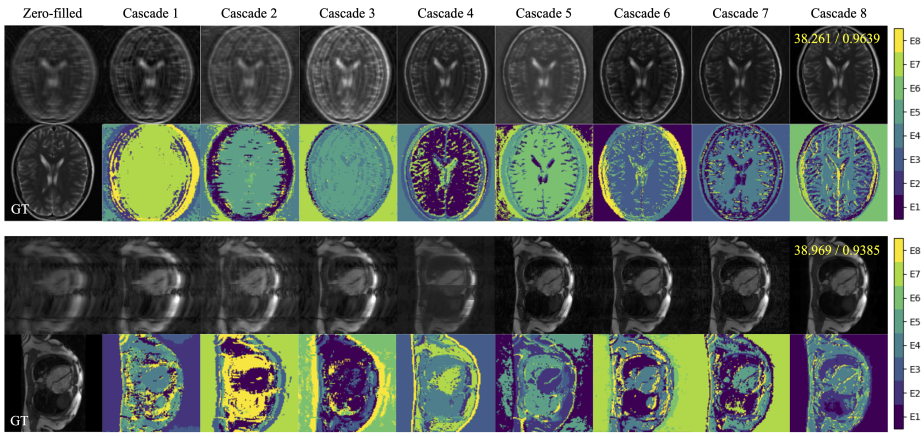

Fig. 2. Visualization of per-cascade reconstruction and Top-1 expert allocation map.

Table 2. Ablation study under 8× acceleration on both datasets. The fastMRI reports configurations (a), (c), and (e), while CMRxRecon reports the full (a)–(e) progression.

	Components			fastMRI 8×			CMRxRecon 8×			Params (M)	
Config.	MoE	Router	Ad. d	PSNR↑	SSIM↑	NMSE↓	PSNR↑	SSIM↑	NMSE↓	Total	Active↓
(a) Dense-B				36.87	0.9134	0.0090	36.29	0.9028	0.0446	109.25	109.25
(b) Dense-L				–	–	–	36.69	0.9044	0.0426	122.29	122.29
(c)	✓	Feature	($d = 0.0$)	36.80	0.9198	0.0092	37.66	0.9141	0.0359	156.32	115.18
(d)	✓	Dual	($d = 0.5$)	–	–	–	37.72	0.9117	0.0344	156.32	115.18
(e) **Ours**	✓	Dual	✓	**37.07**	**0.9228**	**0.0088**	**37.73**	**0.9172**	**0.0339**	156.32	115.18

nounced at 8×. This is consistent with PiMoE's design, as greater ill-posedness leads the physics-informed pathway to assume stronger routing authority, amplifying its contribution to reconstruction quality. Fig. 2 visualizes per-cascade reconstructions alongside the top-1 expert assignment maps from the last routing layer of each cascade, highlighting the distinct specialized roles that emerge across unrolling stages.

Effect of Model Components. Table 2 validates each component through systematic ablation on both datasets under 8× acceleration. The MoE architecture (c) yields superior gains over both the baseline (a) and a parameter-matched dense model (b) on CMRxRecon despite fewer active parameters, confirming the structural efficiency of sparse specialization. However, the benefit of MoE alone is dataset-dependent. In fastMRI, feature-only routing (c) yields mixed results relative to the dense baseline (a), with marginal PSNR degradation despite improvements in SSIM. In contrast, the full PiMoE (e) achieves consistent gains across all three metrics. This pattern aligns with the difference in acquisition conditions between the two datasets. CMRxRecon employs a standardized coil setup where feature-only routing already provides substantial gains, while fastMRI's variable coil configurations present a more challenging routing envi-

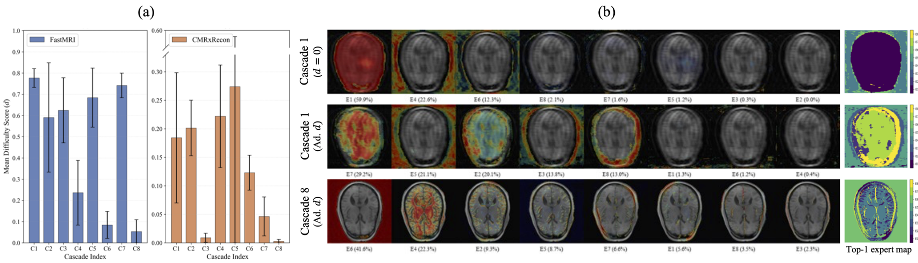

Fig. 3. Dynamics of expert routing across cascades. (a) Mean difficulty score d per cascade. (b) Per-expert top-k routing weight maps, with Top-1 allocation (right) under feature-only routing (d=0, top) and adaptive routing (Ad. d, middle and bottom). Percentages denote token utilization. See Fig. 1 for corresponding descriptors.

ronment that benefits from explicit physics guidance. PiMoE's adaptive difficulty scorer resolves this discrepancy, automatically elevating physics authority when acquisition conditions demand it, achieving consistent improvements across both regimes. The full model (e) recovers both PSNR and SSIM in CMRxRecon, relative to fixed-weight integration (d). Notably, PiMoE achieves these gains with 115M active parameters, comparable to the dense baseline, while the additional expert capacity is only invoked through sparse selection.

Expert Specialization Analysis. Fig. 3(a) depicts the cascade-wise evolution of d. Initially, d averages 0.77 for fastMRI and 0.18 for CMRxRecon, reflecting fastMRI's greater physical ill-posedness from variable coil counts (2–24). Across cascades, both values converge toward 0.09 and 0.03, signaling a transition from physics-driven regularization to anatomical refinement. Fig. 3(b) validates this through direct comparison with feature-only routing. Under d=0, a single expert dominates with nearly 60% utilization. Adaptive routing instead distributes tokens across experts from early cascades, with individual experts attending to regions of high reconstruction difficulty (*e.g.*, E7 for aliasing-affected, E2 for regions with limited coil diversity), while later cascades shift toward anatomical roles (*e.g.*, E6 for white matter, E4 for gray matter). This spontaneous progression confirms that the d effectively orchestrates a two-phase restoration strategy.

Generalization to Unseen Modalities and Views. Through dual routing, PiMoE achieves per-cascade expert specialization as described above, enabling stable operation on unseen modalities and views. To verify this, we evaluated fastMRI T1-Post (n=287) and CMRxRecon Aorta-Tra (n=97) under 8× acceleration. PiMoE outperformed HUMUSNet by +1.44 dB and +1.20 dB in PSNR, and by +0.0094 and +0.0121 in SSIM, respectively, confirming robust generalization to these unseen modalities and views.

4 Conclusion

We presented PiMoE, a framework that addresses the spatially heterogeneous ill-posedness of accelerated MRI reconstruction by replacing monolithic priors with physics-guided expert ensembles. PiMoE's adaptive routing enables spontaneous expert specialization that transitions from physical artifact suppression to anatomical refinement across cascades, establishing an interpretable and robust restoration. An extension is to generalize the physical descriptors beyond Cartesian sampling to non-Cartesian trajectories, broadening its applicability to diverse acquisition protocols.

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

References

1. Fabian, Z., Tinaz, B., Soltanolkotabi, M.: Humus-net: Hybrid unrolled multi-scale network architecture for accelerated mri reconstruction. Advances in Neural Information Processing Systems **35**, 25306–25319 (2022)
2. Hamilton, J., Franson, D., Seiberlich, N.: Recent advances in parallel imaging for mri. Progress in nuclear magnetic resonance spectroscopy **101**, 71–95 (2017)
3. Hammernik, K., Klatzer, T., Kobler, E., Recht, M.P., Sodickson, D.K., Pock, T., Knoll, F.: Learning a variational network for reconstruction of accelerated mri data. Magnetic resonance in medicine **79**(6), 3055–3071 (2018)
4. Huang, Q., Yang, D., Wu, P., Qu, H., Yi, J., Metaxas, D.: Mri reconstruction via cascaded channel-wise attention network. In: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019). pp. 1622–1626. IEEE (2019)
5. Huang, Q., An, Z., Zhuang, N., Tao, M., Zhang, C., Jin, Y., Xu, K., Chen, L., Huang, S., Feng, Y.: Harder task needs more experts: Dynamic routing in moe models. In: Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). pp. 12883–12895 (2024)
6. Huang, S., Li, J., Mei, L., Zhang, T., Chen, Z., Dong, Y., Dong, L., Liu, S., Lyu, M.: Accurate multi-contrast mri super-resolution via a dual cross-attention transformer network. In: Medical Image Computing and Computer Assisted Intervention — MICCAI 2023. p. 313—322. Springer Nature Switzerland (2023)
7. Knoll, F., Clason, C., Diwoky, C., Stollberger, R.: Adapted random sampling patterns for accelerated mri. Magnetic resonance materials in physics, biology and medicine **24**(1), 43–50 (2011)
8. Li, B., Shen, Y., Yang, J., Wang, Y., Ren, J., Che, T., Zhang, J., Liu, Z.: Sparse mixture-of-experts are domain generalizable learners. arXiv preprint arXiv:2206.04046 (2022)
9. Lustig, M., Donoho, D., Pauly, J.M.: Sparse mri: The application of compressed sensing for rapid mr imaging. Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine **58**(6), 1182–1195 (2007)
10. Moratal, D., Vallés-Luch, A., Martí-Bonmatí, L., Brummer, M.E.: k-space tutorial: an mri educational tool for a better understanding of k-space. Biomedical imaging and intervention journal **4**(1), e15 (2008)
11. Pruessmann, K.P., Weiger, M., Scheidegger, M.B., Boesiger, P.: Sense: sensitivity encoding for fast mri. Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine **42**(5), 952–962 (1999)

12. Ren, Y., Li, X., Li, B., Wang, X., Guo, M., Zhao, S., Zhang, L., Chen, Z.: Moe-diffir: Task-customized diffusion priors for universal compressed image restoration. In: European Conference on Computer Vision. pp. 116–134. Springer (2024)
13. Riquelme, C., Puigcerver, J., Mustafa, B., Neumann, M., Jenatton, R., Susano Pinto, A., Keysers, D., Houlsby, N.: Scaling vision with sparse mixture of experts. Advances in Neural Information Processing Systems **34**, 8583–8595 (2021)
14. Robson, P.M., Grant, A.K., Madhuranthakam, A.J., Lattanzi, R., Sodickson, D.K., McKenzie, C.A.: Comprehensive quantification of signal-to-noise ratio and g-factor for image-based and k-space-based parallel imaging reconstructions. Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine **60**(4), 895–907 (2008)
15. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer Assisted Intervention — MICCAI 2015. p. 234—241. Springer Nature Switzerland (2015)
16. Shazeer, N., Mirhoseini, A., Maziarz, K., Davis, A., Le, Q., Hinton, G., Dean, J.: Outrageously large neural networks: The sparsely-gated mixture-of-experts layer. arXiv preprint arXiv:1701.06538 (2017)
17. Sriram, A., Zbontar, J., Murrell, T., Defazio, A., Zitnick, C.L., Yakubova, N., Knoll, F., Johnson, P.: End-to-end variational networks for accelerated mri reconstruction. In: Medical Image Computing and Computer Assisted Intervention — MICCAI 2020. p. 64—73. Springer Nature Switzerland (2020)
18. Tsai, C.M., Nishimura, D.G.: Reduced aliasing artifacts using variable-density k-space sampling trajectories. Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine **43**(3), 452–458 (2000)
19. Wang, Z., Wang, F., Qin, C., Lyu, J., Ouyang, C., Wang, S., Li, Y., Yu, M., Zhang, H., Guo, K., et al.: Cmrxrecon2024: A multimodality, multiview k-space dataset boosting universal machine learning for accelerated cardiac mri. Radiology: Artificial Intelligence **7**(2), e240443 (2025)
20. Xin, B., Ye, M., Axel, L., Metaxas, D.N.: Fill the k-space and refine the image: Prompting for dynamic and multi-contrast mri reconstruction. In: International Workshop on Statistical Atlases and Computational Models of the Heart. pp. 261–273. Springer (2023)
21. 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: proceedings of Medical Image Computing and Computer Assisted Intervention – MICCAI 2024. vol. LNCS 15007. Springer Nature Switzerland (October 2024)
22. Zamfir, E., Wu, Z., Mehta, N., Tan, Y., Paudel, D.P., Zhang, Y., Timofte, R.: Complexity experts are task-discriminative learners for any image restoration. In: Proceedings of the Computer Vision and Pattern Recognition Conference. pp. 12753–12763 (2025)
23. Zbontar, J., Knoll, F., Sriram, A., Murrell, T., Huang, Z., Muckley, M.J., Defazio, A., Stern, R., Johnson, P., Bruno, M., et al.: fastmri: An open dataset and benchmarks for accelerated mri. arXiv preprint arXiv:1811.08839 (2018)
24. Zhang, T., Pauly, J.M., Vasanawala, S.S., Lustig, M.: Coil compression for accelerated imaging with cartesian sampling. Magnetic resonance in medicine **69**(2), 571–582 (2013)
25. Zhang, Z., Romero, A., Muckley, M.J., Vincent, P., Yang, L., Drozdzal, M.: Reducing uncertainty in undersampled mri reconstruction with active acquisition. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 2049–2058 (2019)